

(22) International Filing Date:

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

31 December 1997 (31.12.97)

(30) Priority Data:
08/778,645
60/080,222
3 January 1997 (03.01.97)
US
08/080,222
US

(71) Applicant (for all designated States except US): UNIVER-SITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventor; and
 (75) Inventor/Applicant (for US only): LINK. Christopher [US/US]; 1837 Spruce Street, Boulder, CO 80302 (US).

(74) Agents: MONROY, Gladys, II. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR. BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SI, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: BETA-AMYLOID TOXICITY

(57) Abstract

The present invention relates to methods and compositions for visualization of the toxic effects of transgenes in vivo. In particular, the present invention provides methods and compositions for the production and use of transgenic, including dually transgenic, Caenorhabditis elegans for visualization of the toxic effects of β -amyloid accumulation in vivo.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL AM AT AU AZ BA BB BB BF BG BJ BR CCA CCF CG CII CI CM CN CCC DE	Albania Armenia Austria Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin Brazil Belarus Canada Central African Republic Congo Switzerland Cote d'Ivoire Cameroon China Cuba Czech Republic Germany	ES FI FR GA GB GB GR HU IE II IS IT JP KE KG KP LC LI	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Italy Japan Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechenstein	LS LT LU LV MC MD MG MK MI MN MN MN NE NL NO NZ PL PT RO RU SP	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mauritania Malawi Mexico Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Federation Sudan Sweden	SI SK SN SZ TD TG TJ TM TR TT UA UG US UZ VN YU ZW	Slovenia Slovakia Senegal Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago Ukraine Uganda United States of America Uzbekistan Viet Nam Yugoslavia Zimbabwe
DK EE	Denmark Estonia	LK LR	Sri Lanka Liberia	SE SG	Sweden Singapore		

BETA-AMYLOID TOXICITY

FIELD OF THE INVENTION

The present invention is in the field of medicine and molecular biology. In particular, the invention provides transgenic non-human animals in which expression of a reporter gene is induced by a toxic transgene. The invention also provides methods and compositions for *in vivo* visualization of the toxicity associated with toxic proteins. In addition, the present invention provides methods and compositions for the production and use of transgenic non-human animals for screening candidate drugs to assess and treat toxicity.

10

15

20

25

5

BACKGROUND OF THE INVENTION

Investigation of the molecular mechanisms of various diseases has suggested that many pathologies are caused by expression of proteins that are immediately or gradually deleterious to the cells where they are expressed. Among these diseases are Alzheimer's disease, prion diseases, Huntington's disease, and amyotrophic lateral sclerosis. Several of these diseases are thought to result from aberrant folding of proteins, which results in the accumulation of toxic proteins or protein aggregates.

Alzheimer's disease is associated with the majority of dementia cases in the United States, with an estimated 2 million people afflicted with the disease, and a mortality rate of approximately 100,000 people per year (See, R.W.P. Cutler, "Degenerative and Hereditary Diseases." in Medicine, Scientific American, New York, (1988), pages 11 (IV):1-13; and R. Katzman (1986) N. Engl. J. Med. 314:964). It has been estimated that the total cost for nursing home care alone of Alzheimer's patients exceeds \$13 billion/year (See, M.M. Heckler (1985) Am. Psychol. 40:1240). According to the Centers for Disease Control (CDC), mortality due to Alzheimer's disease in the United States increased 10-fold between 1979 and 1987 (0.4 per 100.000 to 4.2 per 100.000) (See, "Reported death rate for Alzheimer's is up tenfold since 1979," (1990) Clin. Psychiatr. News 18:21).

Patients suffering from Alzheimer's disease typically suffer progressive memory deficit, progressive decline in cognitive functions, anxiety, depression, visuospatial and speech deficits, delusions, personality changes, motor skill deterioration, loss of verbal

30

ability, and incontinence. Eventually, patients are completely incapacitated and disoriented, requiring total care. The course of Alzheimer's disease ranges from less than three years to over 20 years before death occurs. However, in typical cases, it progresses at a fairly constant rate, with an average duration of 6 to 10 years.

5

10

Alzheimer's disease is not the only cause of dementia. Indeed, there are over fifty recognized causes of dementia. As some causes of dementia are amenable to treatment, differential diagnosis of patients suffering dementia is particularly important. Because at least 20% of clinically diagnosed patients were found at autopsy to have had conditions other than Alzheimer's disease, the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and Alzheimer's Disease and Related Disorders Association (ADRDA) retined the clinical diagnostic criteria for Alzheimer's disease (*See.* M.A. Jenike. "Psychiatry." *in Medicine*, Scientific American, New York, [1991], pages 13 (V):1-5). Based on these criteria, the diagnosis of Alzheimer's disease may be "definite" (*i.e.*, requiring examination of brain tissue), "probable" (*i.e.*, patients have deficits in two or more areas of cognition, insidious onset of disease, progressive worsening of memory and other cognitive functions, and normal consciousness levels), or "possible" (*i.e.*, patients meet the criteria for probable Alzheimer's disease, but exhibit variations in the disease course or have a systemic illness that is sufficient to cause dementia, but is not considered to be the cause of the dementia).

20

15

The numerous varieties of dementia, and variations in patient presentations, often make diagnosis problematic. Thus, the NINCDS/ADRDA criteria are very detailed, and necessitate the thorough examination of patients with suspected Alzheimer's disease. Currently, the only way to obtain a definite diagnosis is by post-mortem histological examination of brain tissue for the presence of senile plaques.

25

The histopathological lesions of Alzheimer's disease include neuritic or senile plaques, neurofibrillary degeneration, and granulovacuolar neuronal degeneration. The senile plaques usually contain a core of insoluble, amyloidic extracellular material (" β -amyloid) surrounded by a halo of neurofibrillary tangles and dystrophic neurons. The primary protein component of the amyloidic core of senile plaques is a 4.2 kd amyloid β peptide (often referred to as " $A\beta$ "). A number of other proteins have also been identified

30

as components of senile plaques, including anti-chymotrypsin and apolipoprotein E. The major component of the neurofibrillary tangles is an abnormally phosphorylated microtubule-associated protein referred to as "tau." The β-amyloid protein in senile plaques is a small glycoprotein (*i.e.*. a 39-43 amino acid protein, derived from the larger amyloid precursor protein), that has been detected in non-neural tissues (*e.g.*, skin, subcutaneous tissue, and intestines) and blood vessels of Alzheimer's disease patients (*See*, Cutler, *supra*). Deposits of amyloid may be detected by their ability to bind specific dyes, such as Congo red or thioflavin S. Thus, it has been further hypothesized that detection of this protein may serve as a potential diagnostic aid in the assessment of patients suffering from dementia.

5

10

15

20

25

30

Alzheimer's disease usually arises spontaneously, although genetics play a role in development of the disease. For example, the histopathologic lesions of Alzheimer's disease have been regularly observed in the brains of older patients dying of Down's syndrome (*See*, Cutler, *supra*). It is known that the β-amyloid gene resides on chromosome 21. As Down's patients have an extra copy of this chromosome, there is an increased expression of this chromosome in these patients. It has been hypothesized that increased expression of the β-amyloid gene may regulate the formation of amyloid plaques ("senile plaques") in these patients. In addition, four loci were recently identified as playing a role in the genetic susceptibility of Alzheimer's disease (*See*, Pericak-Vance and Haines (1995) *Trends Genet*. 11:504).

Other potential risk factors for the development of Alzheimer's disease include environmental factors (e.g., head trauma, smoking, and exposure to heavy metals), sociological factors (e.g., depression and educational level), biological factors (e.g., increasing age and hyperthyroidism), and a family history of Alzheimer's disease, Down's syndrome, or Parkinson's disease (See. Pericak-Vance and Haines, supra). Nonetheless, despite recent advances, the exact etiology and pathogenesis of Alzheimer's disease remain largely unknown.

In addition to the problems associated with diagnosing Alzheimer's disease. improvements are needed in the area of treatment. Many agents have been tested for their ability to treat the cognitive decline associated with Alzheimer's disease. For example,

various cholinergic enhancers (e.g., choline and lecithin) have been tested. Unfortunately, cholinergic precursors have been shown to be not useful, although some drugs that stimulate cholinergic transmission may be helpful in some patients. One example is physostigmine, a compound that prevents the synaptic breakdown of acetylcholine. However, the overall clinical effect of this drug has not been as dramatic as initially hoped (See, M.A. Jenike, "Psychiatry," in Medicine, Scientific American, New York, (1991), pages 13 (V):1-5).

Other drugs, such as tetrahydroaminoacridine (THA or tacrine), a centrally acting anti-cholinesterase, have been tested. In a large multi-center trial of THA, liver enzyme abnormalities were reported and the preliminary results indicated that, at least at low dosages. THA is not an effective treatment of Alzheimer's disease (Jenike, *supra*). Ergoloid mesylates (Hydergine) is an extremely safe compound, and remains the most commonly prescribed drug for patients with Alzheimer's disease. However, the overall effects of the drug are at best minimal.

The mechanisms of toxicity due to accumulation of β-peptide are currently the subject of much investigation, and no definitive causes of toxicity have yet been established. (See, Benzi and Moretti (1995) Neurobiology of Aging, 16:661-674). Because of this uncertainty, efforts to develop mechanism-based treatment regimens have not been possible. Compounds presently in use to treat Alzheimer's disease only serve to alleviate the systemic effects associated with the disease.

In order to study disease mechanisms and genetic-based phenomena, animals in which a foreign gene has been inserted have been described by various researchers. International Patent Application WO 96/03034 describes insertion of retroviral vectors into fish, in order to produce fish with desirable traits or to study development. Various transgenic animal models for Alzheimer's disease are described in International Patent Applications WO 93/14200, WO 93/02189, WO 94/12627, WO 94/23049, and European Patent Publication EP 653154. Typically these transgenic animals are mice or other mammals; however, β-amyloid peptide has been expressed in the nematode *Caenorhabditis elegans* (Link (1995) *Proc. Natl. Acad. Sci.* 92: 9368). The use of *C. elegans* for investigation of mutant forms of the *C. elegans* genes mec-4 and deg-1, which

25

5

10

15

20

1.

PCT/US97/23819 WO 98/28971

5

10

15

20

25

30

cause neurodegeneration, is described in U.S. Patent No. 5,196,333. Drawbacks to mammalian animals as model systems are the relatively long generation time, which makes mammals less desirable for high-throughput screening of potential pharmaceuticals, and the difficulty in studying the molecular processes of interest without sacrificing the animal for cell and tissue analysis.

Reporter genes are genes that encode proteins or other compounds that can be detected by a variety of methods, and which "report" the occurrence of successful introduction and expression of gene sequences. β-galactosidase and luciferase are examples of such reporter genes. Recently, the gene for green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* has been described in U.S. Patent No. 5,491,084. This gene can provide a method for indicating expression of a gene via fluorescent detection of GFP. The drawback to using GFP in mammalian systems is as described above, that is, the need to sacrifice the mammal in order to analyze the cells and tissues of interest. This drawback is especially severe in the context of large-scale screening of potential therapeutic compounds, and U.S. Patent No. 5,491,084 does not describe a system suitable for high-throughput screening of pharmaceuticals for activity against toxic proteins expressed in cells.

Thus, methods and animal systems are needed to screen drugs quickly and inexpensively for their effects on proteins and other substances associated with Alzheimer's disease. Convenient methods and animal systems for screening drugs for other neurodegenerative diseases, such as prion diseases. Huntington's disease, and amyotrophic lateral sclerosis, are also desirable.

SUMMARY OF THE INVENTION

The present invention provides methods and compositions useful for the production and use of dually transgenic animals, in particular *Caenorhabditis elegans*. In addition, in one embodiment, the present invention provides transparent animals that express a reporter gene inducible by a toxic transgene. In a preferred embodiment, the toxic transgene encodes β -peptide. Although it is not intended that the present invention be limited to any particular reporter, in a preferred embodiment, the reporter gene is green fluorescence

protein (GFP). It is contemplated that the effects or product of the reporter gene be observable (*i.e.*, the reporter gene is expressed). For example, in embodiments in which GFP is the reporter, its presence is detected using fluorescence microscopy. In addition, in these embodiments, the animals do not need to be sacrificed in order to observe the expression of the reporter gene (*i.e.*, the detection may be accomplished on living animals). If other reporters are used, other detection methods may be necessary. For example, lacZ expression may be detected by exposing the tissues of the animal to the substrate for the gene (*i.e.*, β -galactoside), and observing for the presence of blue dye in the tissues. However, this method requires that the animal be sacrificed in order to observe the expression of the reporter gene.

5

10

15

20

25

- ...- :

In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

In a preferred embodiment, the transparent animal is selected from the class Nematoda. Although it is not intended that the transparent animal of the present invention be limited to any specific animal, in a particularly preferred embodiment, the transparent animal is *Caenorhabditis elegans*.

PCT/US97/23819 WO 98/28971

5

10

15

20

25

30

In an alternate embodiment the cells of the transparent animal display toxicity resulting from the accumulation of β-peptide within the cells. In one preferred embodiment, the genome of the transparent animal comprises SEQ ID NO:5 and SEQ ID NO:8.

The present invention also provides methods for producing dually transgenic non-human animal comprising: providing: a first and second non-human animal; a first transgene comprising β -peptide: and a second transgene, comprising a reporter: introducing the first transgene into the genome of the first non-human animal to produce a first transgenic animal, and introducing the second transgene into the genome of the second non-human animal to produce a second transgenic animal; and mating the first transgenic animal with the second transgenic animal to produce a dually transgenic animal, wherein the β -peptide and the reporter are expressed.

In one embodiment of the methods of the present invention the dually transgenic non-human animal is transparent. In a preferred embodiment, the transparent animal is a nematode, while in a particularly preferred embodiment, the animal is *Caenorhabditis* elegans.

In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

The present invention also comprises methods for testing compounds for anti-toxic effects, comprising: providing a dually transgenic non-human animal expressing a toxic transgene and a reporter; a composition comprising a test compound in a form suitable for administration such that the compound is bioavailable in the cells of the animal; and administering the test compound to the non-human animal. The toxic transgene induces

5

10

15

. 20

25

30

. ...

expression of the reporter gene. In a preferred embodiment, the toxic transgene is β -peptide and the reporter gene is GFP. In one embodiment, the method further comprises the step of measuring a reduction or increase in the expression of the reporter by the dually transgenic non-human animal and thereby identifying a compound as therapeutic. In a particularly preferred embodiment of the methods, the compounds inactivate the β -peptide expressed by the dually transgenic animal.

In one embodiment of the methods for testing compounds for β-peptide toxicity, the dually transgenic non-human animal is transparent. In a preferred embodiment, the transparent animal is a nematode, while in a particularly preferred embodiment, the animal is *Caenorhabditis elegans*. In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is *rol-6(su-1006)*, which produces effects described herein.

In one alternative embodiment, one or more test compounds are tested for their ability to counter the toxic effects of transgene product. It is also contemplated that the test compounds will be tested for their ability to prevent the expression of the toxic transgene, for example, β -peptide.

In alternative embodiments, dually transgenic animals in which the expression of, or effects of transgene toxicity, are reduced or eliminated by the test compounds, are mated. The progeny of these matings are also then tested for the effects of test compounds on the expression of, or effects of toxicity. In yet another embodiment, the progeny of these matings are used in other assay systems for the identification of therapeutic compounds.

In yet another alternative embodiment, the dually transgenic animals of the present invention are used to identify methods suitable for the diagnostic testing of Alzheimer's disease or pathology due to other diseases. Thus, the dually transgenic animals of the present invention are used to develop assays suitable for use in humans or animal models of Alzheimer's disease.

DESCRIPTION OF THE FIGURES

Figure 1 is a diagram illustrating the construction of dual transgenic animals expressing β -amyloid peptide and an hsp/GFP reporter.

Figure 2 shows the DNA sequence (SEQ ID NO:5) and restriction map for pCL12.

Figure 3 is a graphic map of pCL12.

Figure 4 shows the DNA (SEQ ID NO:6), and amino acid sequence (SEQ ID NO:7) of pCL12 from nucleotide 1071 through 1253 (*i.e.*, the β -(1-42) nucleic acid and amino acid sequence).

Figure 5 shows the DNA sequence (SEQ ID NO:8) and restriction map for pGFP-TT.

Figure 6 is a graphic map of pGFP-TT.

5

15

20

25

30

Figure 7 shows the DNA sequence (SEQ ID NO:9) and restriction map for pCL25.

Figure 8 is a graphic map of pCL25.

Figure 9 shows the DNA sequence (SEQ ID NO:10) of rol-6.

Figure 10 shows the amino acid sequence (SEQ ID NO:11) of rol-6.

DESCRIPTION OF THE INVENTION

The present invention provides methods and compositions useful for the production and use of transgenic animals. The methods and transgenic animals of the invention also provide an efficient and effective system for screening drug effective in ameloriating the effects of toxic gene products.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, cell biology and recombinant DNA, which are within the skill of the art. See, e.g., Sambrook, Fritsch, and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, (F.M. Ausubel et al. eds., 1987); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.); PCR 2: A PRACTICAL APPROACH (M.J. McPherson, B.D. Hames and G.R. Taylor eds., 1995); ANIMAL CELL CULTURE (R.I.

Freshney, Ed., 1987); and ANTIBODIES: A LABORATORY MANUAL (Harlow et al. eds., 1987).

Animals

5

10

15

20

25

30

The methods and systems described herein can be practiced with any non-human animal. In a preferred embodiment, the animal used is transparent for at least part of its life cycle, for example, oocytes (*Xenopus* or others); larvae; pupae, fish (zebrafish or others); and nematodes, such as *C. elegans*. *C. elegans* is particularly preferred because it is transparent throughout its life and has been extensively studied as a model to determine developmentally regulated gene expression, as well as pattern formation. The visibility of all of the animals' cells throughout their life cycles, has led to the complete determination of the cell lineages, and a detailed description of the morphogenic changes that occur during *C. elegans* development. Rapid methods for introduction of cloned DNA into the *C. elegans* germ line has provided means to study molecular function and expression *in vivo*, as the effects of the gene expression are usually readily visible in the transparent worms.

Reporter Molecules

The methods and compositions of the present invention make use of reporter genes to monitor gene activity. Visible and quantifiable reporter genes are known and described in the art. Successfully used reporter molecules in gene fusion vectors in studies with *C. elegans* include the *E. coli lac*Z coding region (*See e.g.*, A. Fire *et al.*, (1990) *Gene* 93:189-198), and the *Aequora victoria gfp* coding region (*See e.g.*, M. Chalfie *et al.*, (1994) *Science* 263:802-805; D.C. Prasher *et al.*, (1992) *Gene* 111:229-233; and Genbank Accession #M62654), which produces an intrinsically fluorescent protein. Although various fusion expression vectors have been used and reported in the literature, problems have been encountered (*See e.g.*, A. Fire, "Fire Lab Vector Kit"--June 1995). For example, with *lac*Z, there have been many reports in which it was not possible to correlate transgene expression patterns with physiological expectations, or only weak correlations have been possible. Ectopic expression is often frequently seen with short promoters, and occurs

PCT/US97/23819 WO 98/28971

most prominently in the gut and pharynx of the affected animals. This may be due to weak promoter and/or enhancer signals in the vectors. Expression pattern deficits of *lacZ* fusions have been classified into three groups. In the first, transgenes are expressed in the correct tissue, but mosaic expression is observed (*i.e.*, only a subset of the cells stain during the detection methods). This has been observed even with integrated high-copy transgenes. In the second group, expression in a single tissue or cell population is not seen with the transgene. Finally, there has been the failure of transgene fusion constructs to show expression in the pre-12 cell embryo, or in any embryonic, larval, or adult germ line.

The g/p fusion vectors were developed as an alternative to the lacZ markers, but have been available for a much shorter time (See e.g., Chalfie et al., (1994), supra; U.S. Patent No. 5.491,084). Initial reports indicate that the fluorescence pattern appears to be more restricted than that exhibited by equivalent lacZ fusion. In addition, there are still problems associated with the germline expression of the transgenes and ectopic expression. Indeed, in some cases, the problems with ectopic expression have been exacerbated by these "improvements."

The present invention overcomes many of these problems described in the art. The transparent animals and dually transgenic animals described herein provide easily visible, stably expressed systems in which the toxic transgene and reporter gene are expressed similarly to endogenous chromosomal gene expression. These animals provide an excellent system for screening compounds having effects on the toxic transgene. The present invention, therefore, provides model systems for the study of human diseases and methods of identifying therapeutic compounds using these animal systems.

Inducible Promoters

5

10

15

20

25

30

As discussed above, the reporter genes are operably linked to an inducible promoter. The promoter is induced by the toxic transgene, for example by the gene product of the toxic transgene. It is also contemplated that the toxic transgene can act to induce the promoter indirectly, for example by disrupting other cellular proteins or functions. Suitable inducible promoters are available and can be readily determined by those skilled in the art. Non-limiting examples of promoters which are induced by

"stresses" include the metallothionein gene promoter (*mtl*-1 or *mtl*-2; *e.g.*, Genbank Accession #M92910, #M11794, #X00504, and #X00953), and the *C. elegans* amyloid precursor protein (APP) homolog *apl*-1 gene promoter (*See e.g.*, Daigle and Li, Proc. Natl. Acad., Sci., 90:12045-12049 [1993]). Other potential promoters include those from other inducible heat-shock genes; at least one of the known *C. elegans hsp*70 genes is strongly heat-inducible. It is also contemplated that promoters from genes known to be upregulated under stress conditions in other systems (*e.g.*, superoxide dismutase, catalase, glutathione reductase, etc.) may also be useful.

5

10

15

20

25

In one embodiment, as described in the Examples below, a strong muscle-specific promoter was used to express a potentially secretable form of the β -peptide, so as to generate significant extracellular levels of β -peptide. This was accomplished in order to mimic the situation that may exist in the human brain, and allow observations of cell-external neurotoxicity. The present inventor has also demonstrated that the promoter must be chosen so that enough β -peptide is produced to cause physiological effects, but not so much to kill the animals.

The unc-54/ β -(1-42) minigene was constructed with a modified signal sequence that has been previously shown to allow secretion of a her-1 protein product that is ectopically expressed in muscle cells (M.D. Perry $et\ al.$. (1993) Genes Dev.. 7:216-228). The development of C elegans transgenic for expression of β -amyloid was previously described by the inventor (Link, (1995) Proc. Natl. Acad. Sci., 92:9368-9572).

However, in early experiments, it was observed that β -peptide deposits were not convincingly detected outside of the muscle cells, when tested with the antibodies described in Example 4, below. While an understanding of the mechanism is not necessary for the practice of the present invention, it is apparent that the majority of β -peptide expressed by the transgenic worms is retained in the muscle cells and is responsible for the pathology observed in the muscle cells.

It is also contemplated that β -peptide expression will be directed to other tissues, through utilization of appropriate promoters. For example, it is contemplated that animals expressing β -peptide in the intestine may be particularly useful to analyze compounds such

as drugs for their effect on β -peptide, as these cells readily take up exogenous compounds administered orally.

Transgenes

The transgene element of the present invention can be any sequence which is able to induce the promoter operably linked to the reporter gene. As noted, the transgene is "toxic" in the sense that it disrupts cellular function in some way. Preferably, the transgene encodes a protein that is toxic to the host cells and/or organism in that it causes deleterious effects to the host, for instance interfering with the hosts ability to survive and/or grow. The toxic transgenes encode proteins that disrupt cellular function directly (e.g., the gene product is toxic) or indirectly (e.g., the sequence of the transgene disrupts cellular function by some mechanism other than its gene product). Toxic proteins are distinguishable from simple chemical toxins (e.g., heavy metals and the like) by their

antigenicity and higher molecular weight.

'In one embodiment, the toxic transgene comprises a gene encoding for an amyloidic protein, for example beta amyloid peptide, prion protein variants, transthyretin variants, gelsolin variants, cystatin variants, lysozyme variants and the like. In another embodiment, the transgene encodes a protein containing polyglutamine resulting from triplet-repeat expansion such as huntingtin (a protein that has been implicated in Huntington's Chorea). ataxin-1 or ataxin-2. Alternatively, proteins associated with inherent amyelotopic lateral sclerosis (ALS) for example, superoxide dismutase 1 variants and over-expressed neurofilament protein, can be used. It will be understood that the transgene can encode for an entire toxic protein or, alternatively, a functional (i.e. toxic) fragment.

25

30

20

15

Transgenic Animals

In one aspect, the present invention includes a dually transgenic non-human animal. In one embodiment, the development of these dually transgenic animals involves the production of two *C. elegans* lines with a single transgene in each line, designated as CL2005 and CL2070. *C. elegans* was chosen due to the observation that the toxic effect of

 β -amyloid has been reported to occur in these animals, in addition to humans. Line CL2005 exhibited muscle-specific expression of human β peptide, while line CL2070 exhibited stress-inducible expression of GFP. GFP was incorporated into the present invention as it provides a marker (*i.e.*, "reporter") that is readily visible in living worms. Thus, dually transgenic animals both express β -peptide in their muscle cells and under appropriate conditions for observation (*i.e.*, fluorescence microscopy), exhibit green fluorescence in these cells.

Mating of these lines resulted in the production of dually transgenic animals, in which the presence of β -amyloid was easy to detect, due to the fact that the dually transgenic animals expressing β -amyloid glowed green. Thus, the present invention provides an easily detectable method for the specific expression of β -peptide in dually transgenic animals. This ease of detection provides great advantages for the development of methods to analyze the effects of β -amyloid in vivo. In addition, because the dually transgenic animals are not killed in order to detect the presence of β -amyloid, as would be required if other reporters (e.g., lacZ or luciferase) were used, the animals may be useful for screening compounds for their effects on β -amyloid in an in vivo situation. Thus, the use of an reporter which is detectable in vivo provides significant advantages over currently available methods.

Other Advantages of Transgenic C. elegans

5

10

15

20

25

30

The present invention also includes methods of screening compounds for their ability to prevent or inhibit toxicity due to the toxic transgene. In these drug screening embodiments, the methods and animals described above can be used to identify potential therapies. For example, in assays to develop compounds to block the expression of β -amyloid or counter its toxic effects, transgenic worms (e.g., green fluorescing worms with muscle-specific β -peptide expression), would be observed for their loss of fluorescence. In assays where the compound is effective in countering the expression of β -peptide or effects of β -peptide, the animals would no longer glow green. It is easy to visually screen and quantify (using commercially available equipment) for animals that have lost the ability to glow, and then use them to detail the interaction between the compound and the toxic

WO 98/28971

transgene. The present invention provides a fast and efficient screening system, for PCT/US97/23819 example by using commercially available equipment for assaying multiple compounds at

Thus, the transgenic animals of the present invention expressing a detectable reporter gene triggered by a toxic transgene provide animal models for human diseases. 5 For instance, a reporter gene (such as GFP) induced by expression of β -peptide provides an animal model for human Alzheimer's disease. It is also not intended that the present invention be limited to animals expressing β -peptide in conjunction with GFP. It is contemplated that expression of other proteins associated with dementia and/or

Alzheimer's or other diseases would also find use in conjunction with GFP in C. elegans as 10 well. For example, it is contemplated that proteins (including mutated versions of proteins) such as the presentlins associated with Alzheimer's disease (See e.g., J. Marx (1996) Science 274:1838-1840) will be used in conjunction with GFP and full-length APP in dually transgenic C elegans. Other toxic proteins which provide animal models for 15 other human diseases are described herein.

These transgenic animals also provide a means to screen compounds for their ability to decrease or eliminate the toxicity. Examples of compounds suitable for testing using the transgenic animals of the present invention include, but are not limited to, Congo Red, tumor necrosis factor (TNF), estrogen, tacrine (9-amino-1,2,3,4-tetrahydroacridine), dihydroepiandrosterone (DHEA). compounds that inhibit ApoE4. and others. commercially available from suppliers such as Sigma. Compounds to be tested for antitoxic activity are administered to the same number of dually transgenic animals (e.g., generated using the methods described in the Examples, below) from the control group and the treatment group, and the presence or absence of reporter gene used as a measure of efficacy.

20

25

30

_9828971A2 .

The compounds being tested can be administered using any suitable route (e.g., oral, parenteral, controlled-release transdermal methods, and implants, etc.). In one preferred route, the compounds to be tested are suspended in the growth media provided to the worms. Generally speaking, the route of administration will depend on the stability of the compound, the susceptibility of the compound to "first pass" metabolism, the

concentration needed to achieve a therapeutic effect, and the like. Following initial screening, a compound that appears promising (i.e., which increases the number of worms which display reduced β -peptide toxicity) is further evaluated by administering various concentrations of the compound to additional transgenic animals in order to determine an approximate therapeutic dosing range.

5

10

15

20

25

30

Another screening method involves the crossing of the transgenic worms of the present invention with other transgenic worms. The animals are observed after treatment, in the presence and absence of the test compound(s), with the effects on the toxic transgene being gauged either by crude survival or the presence/absence of the reporter gene. It is also contemplated that the methods of the present invention be modified so as to provide means to analyze disease-related proteins believed to have dominant toxic effects, by substituting appropriate sequences for the β -peptide sequences used as described in the Examples. For example, proteins such as transthyretin (known to be associated with familial amyloid polyneuropathy; See e.g., Christmason et al., (1991) FEBS 281:177-180; Genbank Accession #D00096), and variant superoxide dismutase (known to be associated with familial amyotrophic lateral sclerosis [Lou Gehrig's disease]), prion proteins, A4 amyloid protein (See e.g., Ponte et al., (1988) Nature 331:525-527; Salbaum et al., U.S. Patent No. 5,151,508, herein incorporated by reference), APP (See e.g., Kitaguchi et al., (1988) Nature 331:530-532; Sata et al., EP Appln. 94117512.7; Scott et al., WO 9412627; Wadsworth et al., WO 9314200; Gearhart et al., WO 9423049; and Neve et al., WO 9302189, all of which are herein incorporated by reference), other amyloidic proteins (e.g., variant lysozymes and amylin peptide), and other proteins associated with neuronal degeneration (See e.g., U.S. Patent No. 5,196,333), will be used in the methods of the present invention by substituting the appropriate nucleic acid sequences encoding the protein of interest for the β -peptide described in Example 1. It is also contemplated that animals transgenic for dual proteins (e.g., \beta-peptide in combination with another protein. such as transthyretin) will be used. Thus, the present invention provides the means to analyze the effects of numerous genes and proteins in vivo.

Furthermore, the transgenic animals of the present invention provide distinct advantages over other transgenic animals currently used to analyze diseases such as

Alzheimer's disease. Due to the short gestation period of *C. elegans*, transgenic animals can be produced much more rapidly than when mammals, such as mice are used. For example, transgenic mice overexpressing a 695 amino acid isoform of β-amyloid precursor did not show learning and memory impairment until they were 9-10 months of age (Hsiao *et al.*, Science 274:99-102 [1996]). In contrast, the transgenic animals of the present invention express β-peptide and may be manipulated at a very early age, even in embryonic stages.

5

0

In addition, because the transgenic animals are easy to select based on the presence of the green fluorescence, invasive procedures such as surgery, necessary to analyze the effects of the toxic transgene in other animals (e.g., mice) are avoided.

Finally, the apparent toxicity of β-peptide in transgenic animals was found to be temperature-dependent. Transgenic animals maintained at 25.5°C were significantly more sick than those maintained at 16°C. The animals maintained at the higher temperature became paralyzed more quickly, failed to eat or grow, failed to show normal egg-laying, and many died before reaching adulthood. However, wild-type animals grow well at either of these temperatures. The CL2005 parental line was temperature-sensitive for viability, as it was not possible to propagate this line at the elevated temperature. In the CL2070 parental line, 25.5°C maintenance was not sufficient to induce the hsp/GFP construct. The hsp/GFP response was similarly temperature-dependent. It was difficult to detect GFP when the dual transgenic animals were raised at 16°C, but GFP induction was dramatic when the animals were raised at 25.5°C.

Line of the

These temperature dependency observations were exploited by the propagation of animals at 16°C, and then upshifting them to 25.5°C, when GFP induction was desired. This effect also has potential benefits in the analysis of compounds, as the animals can be pre-incubated in the presence of drugs or other compounds at the lower temperature before the upshift to the higher temperature, in order to ensure that the presumed protective effect of the drug was in place, prior to the strong induction of β-peptide toxicity. Importantly, these assays may be completed within one day and are suitable for rapid methods (c.g., the use of a microtiter format and a plate fluorimeter), so that literally thousands of compounds may be tested simultaneously.

In sum, the present invention provides methods and compositions useful as animal models for disease, as well as providing methods and compositions for disease therapy and prevention. The animal testing may be supplemented and confirmed by testing on human subjects. However, the transgenic animals of the present invention allow the testing of a large number of compounds, both various methods, including those known in the art.

DEFINITIONS

To facilitate understanding of the invention, a number of terms are defined below.

The term "non-human animal" includes vertebrates such as rodents, arthropods, insects (e.g., Diptera), fish (e.g., zebrafish), non-human primates, ovines, bovines, ruminants, lagomorphs, porcines, caprines, equines, canines, felines, aves, etc. The non-human animal can be used at any stage in its development, for example oocyte, fetal, larval, pupal stages or the like. Preferred non-human animals are those that are transparent, such as certain nematodes, oocytes, larvae and fish.

As used herein, the term "transparent" is used in references to animals through which light will be transmitted. However, it is not intended that the amount of light transmittance be limited to any particular amount. For example, an animal is transparent so long as at least some light may be transmited through its body. Non-limiting examples of transparent animals are larval stages of some animals (e.g., flies) and oocytes (e.g., Xenopus oocytes). Particularly preferred non-human animals are selected from the nematodes (i.e., any animal in the Class Nematoda), most preferably Caenorhabditis

elegans. However, it is contemplated that other transparent animals, such as zebrafish will

be useful in the present invention. For example, transgenic zebrafish have been produced

(See e.g., Lin et al., PCT Publ. WO9603034; incorporated herein by reference).

The "non-human animals having a genetically engineered genotype" of the invention are preferably produced by experimental manipulation of the genome of the germline of the non-human animal. These genetically engineered non-human animals may be produced by several methods including the introduction of a "transgene" comprising nucleic acid (usually DNA) into an embryonal target cell or integration into a chromosome of the somatic and/or germ line cells of a non-human animal by way of human

25

5

10

15

20

intervention, such as by the methods described herein. The process by which a DNA molecule becomes stably incorporated into another genome is referred to as "stable integration." Non-human animals which contain a transgene are referred to as "transgenic non-human animals". A transgenic animal is an animal whose genome has been altered by the introduction of one or more transgenes.

5

10

15

20

25

30

)_ 04200-

The term "transgene" as used herein refers to a foreign gene that is placed into an organism by introducing the foreign gene into gonadal cells, embryonic cells, newly gene sequence) which is introduced into the genome of an animal by experimental introduced gene contains some modification relative to the naturally-occurring gene. A manipulation gene is the β -amyloid gene (e.g., β -peptide), or fragments thereof. A for example, the toxic transgene may produce a gene product (i.e., protein) which is toxic transgene. Alternatively, the transgene may disrupt other cellular proteins, or act as a DNA binding site without producing a protein

As used herein, the term "toxicity" refers to the production of toxic effects by any compound or substance. For example, while it is not necessary to the understanding and use of the present invention, β -peptide toxicity may arise due to the accumulation of β -peptide molecules. There are numerous manifestations of toxicity that may occur. For example, the toxicity associated with β -peptide may be manifested as negative effects on muscle cells of dually transgenic animals.

The terms "promoter element," "promoter," or "promoter sequence" as used herein, refer to a DNA sequence that is located at the 5' end of (i.e., precedes) a gene in a DNA polymer and provides a site for initiation of the transcription of the gene into mRNA. An instance binding of a transcription factor, stress heat or the like

The term "reporter gene" as used herein refers to genes that encode proteins or other compounds that can be detected by a variety of methods. These genes "report" the occurrence of successful introduction and expression of sequences such as transgenes.

Non-limiting examples of reporter genes include antibiotic resistance genes, genes encoding enzymes and genes encoding other detectable proteins. Expression of the reporter gene is detected using methods known in the art. In a preferred embodiment, the reporter gene is a GFP gene, although lacZ β -galactosidase gene, or any other reporter system may be used to detect the successful production of transgenic animals. In particularly preferred embodiments, the reporter is a compound or protein which may be present or expressed within living animals. That is, it is not necessary to sacrifice the animal in order to detect the presence of the reporter. The type of the reporter gene system used is not critical to the invention, and it is contemplated that any system suitable for use with the transgenic animals of the present invention will be used.

5

10

15

20

25

30

As used herein, the terms "dual trangenic" and "dually transgenic" refer to animals or cells in which more than one transgene have been introduced. For example, the term is used in reference to cells which contain the sequences encoding β -peptide, and sequences encoding a reporter (e.g., GFP). However, it is not intended that the number of transgenes in the dually transgenic animals of the present invention be limited to two. For example, the transgenic animals may also contain another one or more marker genes (e.g., rol-6 sequences), in addition to β -peptide and reporter sequences. In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

The transgenic animals of the present invention are preferentially generated by introduction of the targeting vectors into gonad cells. Transgenes can be efficiently introduced into the cells by DNA transfection using a variety of methods known to the art, including electroporation, calcium phosphate co-precipitation, protoplast or spheroplast fusion, lipofection, and DEAE-dextran-mediated transfection. Transgenes may also be introduced into cells by retrovirus-mediated transduction or by micro-injection. In one preferred embodiment, the transgenes are injected into gonads of *C. elegans* as described by Mello *et al.* (1991) *EMBO J.*. 10:3959-3970. Alternative methods for the generation of transgenic animals containing an altered gene are known to the art. For example,

embryonal cells at various developmental stages can be used to introduce transgenes for the production of transgenic animals. Different methods are used depending on the stage of development of the embryonal cell.

5

10

15

)

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting related sequences include oligolabeling, nick translation, and end-labeling or PCR amplification using a labeled nucleotide. Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. The term may also be used in reference to proteins. For example, a variety of protocols which employ polyclonal or monoclonal antibodies specific for the β-peptide protein product are known in the art (See, the Examples). These antibodies can be used as markers for the expression of proteins. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS).

The term "gene" refers to a DNA sequence that comprises coding sequences necessary for the production of a polypeptide or precursor. In some instances, a gene can also include control sequences. The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired enzymatic activity is retained.

The term "gene of interest" refers to any gene, the manipulation of which may be deemed desirable for any reason, by one of ordinary skill in the art.

The term "wild-type" refers to a gene or gene product which has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designed the "normal" or "wild-type" form of the gene. In contrast, the term "modified" or "mutant" refers to a gene or gene product which displays modifications in sequence and or functional properties (i.e., altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated: these are

3/28971 PCT/US97/23819

tified by the fact that they have altered characteristics when compared to the wild-type or gene product.

The terms "targeting vector" or "targeting construct" refer to oligonucleotide inces comprising a gene of interest flanked on either side by regulatory sequences. ably, the targeting vector is capable of homologous recombination such that the gene rest is integrated by recombination.

As used herein, the terms "vector" and "vehicle" are used interchangeably in ice to nucleic acid molecules that transfer DNA segment(s) from one cell to another. The terms "expression vector" or "expression cassette" as used herein, refers to a inant DNA molecule containing a desired coding sequence and appropriate nucleic tuences necessary for the expression of the operably linked coding sequence in a ir host organism. Nucleic acid sequences necessary for expression in prokaryotes include a promoter, an operator (optional), and a ribosome binding site, often along it sequences. Eukaryotic cells are known to utilize promoters, enhancers, and on and polyadenylation signals.

ne terms "in operable combination," "in operable order," and "operably linked" as n, refer to the linkage of nucleic acid sequences in such a manner that a nucleic rule capable of directing the transcription of a given gene and/or the synthesis of rotein molecule is produced. The term also refers to the linkage of amino acid n such a manner so that a functional protein is produced.

property with processing

term "tissue-type specific" as it applies to a promoter, refers to a promoter that f selectively directing expression of a gene in a specific tissue. Similarly, the ne-specific promoter in the method of the present invention does not require cificity. In general, the requisite specificity is found where a plurality (or, bly, a majority) of cells in one tissue type express a gene of interest, while e.g., greater than 80%, and preferably greater than 90%, and more preferably 5%), of the cells in other tissue types do not. In one embodiment of the ion, the strong muscle promoter unc54 was used (See, Example 1).

In the term "portion" when in reference to a gene refers to fragments of fragments may range in size from a few nucleotides to the entire gene

SHOCTION

sequence minus one nucleotide. Thus, "an oligonucleotide comprising at least a portion of a gene" may comprise fragments of the gene or the entire gene.

As used herein, the term "substantially purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. An "isolated polynucleotide" is therefore a substantially purified polynucleotide.

5

As used herein, the term "polymerase chain reaction" ("PCR") refers to the method of K.B. Mullis disclosed in U.S. Patent Nos. 4,683,195, 4.683,202 and 4,965,188, all of which are hereby incorporated by reference, which describe a method for increasing the 10 concentration of a segment of a target sequence in a mixture of genomic DNA without cloning or purification. This process for amplifying the target sequence consists of introducing a large excess of two oligonucleotide primers to the DNA mixture containing the desired target sequence, followed by a precise sequence of thermal cycling in the presence of a DNA polymerase. The two primers are complementary to their respective 15 strands of the double stranded target sequence. To effect amplification, the mixture is denatured and the primers then annealed to their complementary sequences within the target molecule. Following annealing, the primers are extended with a polymerase so as to form a new pair of complementary strands. The steps of denaturation, primer annealing and polymerase extension can be repeated many times (i.e., denaturation, annealing and 20 extension constitute one "cycle": there can be numerous "cycles") to obtain a high concentration of an amplified segment of the desired target sequence. The length of the amplified segment of the desired target sequence is determined by the relative positions of the primers with respect to each other, and therefore, this length is a controllable parameter. By virtue of the repeating aspect of the process, the method is referred to as the 25 "polymerase chain reaction" (hereinafter "PCR"). Because the desired amplified segments of the target sequence become the predominant sequences (in terms of concentration) in the mixture, they are said to be "PCR amplified."

With PCR, it is possible to amplify a single copy of a specific target sequence in genomic DNA to a level detectable by several different methodologies (e.g., hybridization

with a labeled probe; incorporation of biotinylated primers followed by avidin-enzyme conjugate detection; and/or incorporation of ³²P-labeled deoxynucleotide triphosphates, such as dCTP or dATP, into the amplified segment). In addition to genomic DNA, any oligonucleotide sequence can be amplified with the appropriate set of primer molecules. In particular, the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications. Amplified target sequences may be used to obtain segments of DNA (e.g., genes) for the construction of targeting vectors, transgenes, etc.

5

10

15

20

25

30

As used herein, the terms "PCR product" and "amplification product" refer to the resultant mixture of compounds after two or more cycles of the PCR steps of denaturation, annealing and extension are complete. These terms encompass the case where there has been amplification of one or more segments of one or more target sequences.

The term "Central Nervous System" refers to the "spinal cord" and the "brain." The spinal cord comprises white areas and grey areas. The grey area contains nerve cell bodies, whereas the white area is essentially comprised of myelinated nerves. The brain, which is also known as the encephalon, is that portion of the cerebrospinal axis which is contained in the cavity of the cranium. The brain comprises the two cerebral hemispheres, the interbrain, the mid-brain, the pons Varolli and cerebellum, and the medulla oblongata. The two hemispheres together with the parts derived from the thalamencephalon form the forebrain. The two cerebral hemispheres are separated by the longitudinal fissure and also comprise the bylvian fissure, the fissure of Rolando, and the parieto-occipital fissure. The lobes on the external surface of the brain comprise the frontal lobe, the parietal lobe, the occipital lobe, and the temporal lobe. Placed along the middle line of the brain are, among others, the rostrum and peduncles of corpus callosum, lamina cinera, optic commissure and the pituitary body. On each side of the middle line lies the frontal lobe, olfactory lobe and the hemisphere of cerebellum.

The terms "neuron," "neural cell," and "nerve cell" are used interchangeably to refer to a cell which is located in the nervous system. Nerve cells are composed of the nerve cell body (perikaryon), one or more dendrites, and an axon. Neurons can be classified according to the number of processes originating from the cell body. Thus,

unipolar neurons have a single process, bipolar neurons have one axon and one dendrite, while multipolar neurons (which are the most common) comprise more than two processes. The term "neuron" comprises cholinergic neurons and sensory neurons. As used herein, the term "cholinergic neuron" means a neuron in the Central Nervous System (CNS) and in the Peripheral Nervous System (PNS) whose neurotransmitter is acetylcholine. As used herein, the term "sensory neuron" includes a neuron which is responsive to environmental cues (e.g., temperature and movement) from, for example, the skin, muscle and joints of a mammal.

5

10

15

20

25

30

The term "nerve" refers to two or more neurons arranged in linear sequence such that the axon of one neuron establishes a structural and functional link with the dendrite of a second neuron to form a "synapse."

The term "compound" refers to any chemical entity, pharmaceutical, drug, and the like that can be used to treat or prevent a disease, illness, sickness, or disorder of bodily function. Compounds comprise both known and potential therapeutic compounds. A compound can be determined to be therapeutic by testing using the testing methods of the present invention (i.e., a "test compound"). A "known therapeutic compound" refers to a therapeutic compound that has been shown (e.g., through animal trials or prior experience with administration to humans) to be effective in such treatment or prevention. In other words, a known therapeutic compound is not limited to a compound efficacious in the treatment or prevention of neural related disorders.

A compound is said to be "in a form suitable for administration such that the compound is bioavailable in the blood of the animal" when the compound may be administered to an animal by any desired route (e.g., oral, intravenous, subcutaneous, intramuscular, etc.) and the compound or its active metabolites appears in the blood of the animal in an active form. Administration of a compound to a pregnant animal may result in delivery of bioavailable compound to the embryonic progeny of the animal.

The "wild-type β-amyloid" or "β-peptide" gene and gene product refers to the nucleotide and amino acid sequences provided in SEQ ID NOS:6 and 7, respectively. Those skilled in the art will be well aware that certain modifications of SEQ ID NOS:6 and 7 can be made which will not interfere with the production of a polypeptide having an

_ - 25 -

activity indistinguishable from that of the wild-type β-amyloid; the present invention specifically contemplates these variant forms of β-amyloid. A "variant" of the β-peptide is defined as an amino acid sequence that differs by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions (i.e., additions), or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNAStar software.

The term "an oligonucleotide sequence comprising at least a portion of a human β -amyloid gene" refers to a polynucleotide sequence (*i.e.*, a nucleic acid sequence) containing a nucleotide sequence derived from a human β -amyloid gene. This sequence may encode a portion or all of the β -amyloid protein; alternatively, this sequence may contain non-coding regions derived from the β -amyloid gene or a combination of coding and non-coding regions. The oligonucleotide may be RNA or DNA and may be of genomic or synthetic origin.

As used herein the term "portion" when in reference to a gene refers to fragments of that gene. The fragments may range in size from 10 nucleotides to the entire gene sequence minus one nucleotide. Thus, "an oligonucleotide comprising at least a portion of a gene" may comprise fragments of the gene or the entire gene.

An animal whose genome comprises a "heterologous marker gene" is an animal whose genome contains a marker gene not naturally found in the animal's genome. In one preferred embodiment of the present invention, the heterologous marker gene is a mutant collagen gene, such as the rol-6(su-1006) gene. However, it is intended that other marker genes will be used with success in the present invention, including other mutant collagen genes, as well as other marker genes commonly known to those in the art.

As used herein, the term "diagnostic assay" refers to methods for the diagnosis of disease, illness, and/or pathology. It is intended that the term encompass any methods for

diagnosis, including, but not limited to assays based on immunoreactivity (e.g., radioimmunoassays, fluorescence immunoassays, enzyme immunoassays), histochemistry, dye retention or binding (e.g., fixing of dyes such as Congo Red), nucleic acid based diagnostic methods (e.g., identification of nucleic acid sequences associated with disease or pathology), etc.

EXPERIMENTAL

The following examples are provided in order to demonstrate and further illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the experimental disclosure which follows, the following abbreviations apply: eq (equivalents); M (Molar); μM (micromolar); N (Normal); mol (moles); mmol (millimoles); μmol (micromoles); nmol (nanomoles); g (grams); mg (milligrams); μg (micrograms); ng (nanograms); l or L (liters); ml (milliliters); µl (microliters); cm (centimeters); mm (millimeters); μm (micrometers); nm (nanometers); °C (degrees Centigrade); U (units), mU (milliunits); min. (minutes); sec. (seconds); % (percent); kb (kilobase); bp (base pair); PCR (polymerase chain reaction); β-(1-42) amino acids 1-42 of β-amyloid peptide; Tris (tris(hydroxymethyl)-aminomethane); BSA (bovine serum albumin); Fisher (Fisher Scientific, Pittsburgh, PA); Sigma (Sigma Chemical Co., St. Louis, MO.); Promega (Promega, Corp., Madison, WI); Perkin-Elmer (Perkin-Elmer/Applied Biosystems, Foster City, CA); Senetek (Senetek, PLC, Maryland Heights, MO); Boehringer Mannheim (Boehringer Mannheim, Corp., Indianapolis, IN); Stratagene (Stratagene, Inc., La Jolla, CA); and NEB (New England Biolabs, Beverly, MA). Unless otherwise indicated, the restriction enzymes used in these Examples were obtained from NEB. C. elegans is available from the Caenorhabditis Genetics Center, at the University of Minnesota, St. Paul, MN.

EXAMPLE 1

Assembly of the β -(1-42) Minigene

5

10

15

20

25

PCT/US97/23819

1-42) minigene used in subsequent experiments (See. ic C. elegans engineered to express amyloidic human escribed below, these animals contained constructs in 54 promoter/enhancer of C. elegans drove the expression of lerived from human β-amyloid cDNA clones, as described ıman B-amyloid peptide in transgenic Caenorhabditis , 92:9368-9372 [1995]). ng the 42 amino acid β-amyloid peptide derived from in cDNA was assembled. The artificial signal peptide 52.81 was amplified under standard conditions, using "GGCCAAAGGACCC-3')(SEQ ID NO:1), and "SP-down" CCAGCAAGAT-3')(SEQ ID NO: 2), cleaved with NheI nd inserted between the unique Nhel and Kpnl sites of : construct "pCL2." This process resulted in a re-2. such that the signal peptide cleavage site, as predicted by von Heijne, Nucl. Acids Res., 14:4683-4690 [1986]), Gly-Thr dipeptide encoded by the KpnI site. ding amino acids 1-42 of β-amyloid, and which contained olified using standard methods, from human \(\beta\)-amyloid 4T4B (P. Ponte et al., Nature 331:525-527 [1988]), by GGGGTACCGATGCAGAATTCCGACATGA-3') (SEO 5'-CCCGAGCTCACGCTATGACAACACCGCCAAcation product was cleaved with KpnI and SacI, and il and Sac I sites of pCL2, to generate "pCL3." 12) minigene fragment was removed from this plasmid by d inserted between the unique Nhel and Sacl sites of The sequence of the β -(1-42) minigene was confirmed by coding strand only, by techniques known in the art. The 0:5), is shown in Figure 2. A graphic map of pCL12 is nows the DNA and amino acid sequences (SEQ ID NOS:

6 and 7, respectively) of pCL12 from nucleotide 1071 through 1253 (i.e., the $\beta(1-42)$ nucleic acid and amino acid sequences) present in the construct.

CA1Ca (AGABA / ABACC 666. 4668A / ABACC 666. 4668A / ABACC

and the second second second

EXAMPLE 2

Construction of the hsp/GFP Reporter

imple, the expression vector ("pCL25") containing the hsp/GFP reporter expression of the GFP-TT gene was constructed. This vector was used in ments (See, Example 3) to produce transgenic C. elegans was 2-TT is a modified form of GFP that contains Ser65Thr and Ile167Thr relative to the sequence of the wild-type GFP). The substitutions present ce a protein which results in the production of a much brighter the wild-type GFP.

as constructed as described below. A 431 bp *Hin*dIII-*Bam*HI fragment p16-2 promoter was excised from the plasmid "pPD49.78." by digestion *Hin*dIII. This 431 bp *Hin*dIII-*Bam*HI fragment was inserted between the de #1) and *Bam*HI (nucleotide #31) sites of the GFP-TT gene in the TT," (available from Yishi Jin. at the University of California, Santa igase (Promega). The pGFP-TT plasmid contains the coding regions for into a Tu61 backbone. The DNA sequence of pGFP-TT (SEQ ID NO: 8) re 5. A schematic map of pGFP-TT is shown in Figure 6. in which unique are shown. The pGFP-TT plasmid was digested with *Hin*dIII and *Bam*HI, ragment containing the hsp16-2 promoter was inserted, in order to ss-inducible expression vector "pCL25." The DNA sequence of pCL25 and restriction map are shown in Figure 7. Figure 8 provides a schematic n. which unique restriction sites are shown.

5

10

15

20

25

30

EXAMPLE 3

Construction of Transgenic Animals

In this Example, the transgenic parent animals were produced, with one line expressing the β(1-42) minigene (designated as "CL2005") and the other line expressing the hsp/GFP reporter (designated as "CL2070"). For both lines, the transgenes were introduced into *C. elegans* by gonad microinjection as known in the art and described by Mello *et al.* (Mello *et* EMBO J., 10:3959-3970 [1991]). Marker plasmid pRF4 containing the gene (rol-6[su1006]) (SEQ ID NO:9)(pRF4 carries a 4 kb EcoRI fragment of *C. elegans* genomic DNA containing the rol-6[su1006] collagen gene in the Bluescribe vector [Stratagene]: See. Mello *et al.*, supra; and Kramer *et al.*. Mol. Cell. Biol., 10:2081-2090 [1990]) was coinjected with the constructs, at approximately 100 ng/μl for each plasmid. Into morphologically wild-type animals, and Roller transgenic progeny were recovered. The (rol-6[su1006]) gene (SEQ ID NO:9), contained within the pRF4 plasmid is a mutated *C. elegans* collagen gene, the expression of which produces the dominant, distinctive "Roller" phenotype. Figures 9 and 10 show the DNA and amino acid sequences of the rol-6 gene used in this Example, respectively.

Transmitting lines were established and maintained by selection for the Roller marker phenotype. Transgenic animals produced in this manner maintain the injected marker phenotype. Transgenic animals produced in this manner maintain the injected DNA as an extrachromosomal, multicopy array of variable mitotic and meiotic stability. Strains containing chromosomally integrated transgenes were recovered by irradiation of lines containing extrachromosomal transgenic arrays with 7000 rad (1 rad = 0.01 Gy) of lines containing extrachromosomal transgenic arrays with 7000 rad (1 rad = 0.01 Gy) of lines were then screen for 100% transmittance of the marker transgene. The transgenes in both the CL2005 and CL2070 lines were chromosomally integrated, and were 100% stable.

Transgenic (i.e., as indicated by the Roller phenotype) animals produce both transgenic (i.e., Roller) and non-transgenic (i.e., non-Roller) progeny. These non-transgenic progeny were found to serve as good internal controls for phenotypic and immunohistochemical comparisons. The expression of GFP in the dual transgenic animals immunohistochemical comparisons. The expression of animals from 16 to 25.5°C, at all can be detected in less than 24 hours after the upshift of animals from 16 to 25.5°C, at all

PCT/US97/23819 WO 98/28971

stages of development, from late embryonic to adult. Expression of GFP resulted in the production of green fluorescence in muscle cells; intense tissue-specific expression of GFP was observed. The fluorescence can be observed using compound or dissecting epifluoroescence microscopy (*i.e.*, with standard fluorescein excitation and emission filters). It is also contemplated that the fluorescence is observable by use of fluorimeters and cell sorters.

EXAMPLE 4

Immunohistochemistry of Transgenic Animals

10

5

In this Example, immunohistochemistry was used to confirm the transmittance of the chimeric constructs in large populations of putative integrated lines.

As described by Link (Link, 1995, supra), whole mount specimens were prepared by fixing animals in 4% paraformaldehyde and permeabilizing them with 2-mercaptoethanol collagenase as known in the art and described by Link et al. (Link et al., Genetics 131:867-881 [1992]). Three antibodies were used (polyclonal rabbit anti-β peptide antibody, Boehringer Mannheim; mouse monoclonal antibody Ab 4G8, available from Senetek: and monoclonal 4.1, a gift from B. Cordell at Scios Nova). The monoclonal 4.1 recognizes residues 8-15 of β-peptide.

20

25

30

15

As was observed by Link (Link, 1995, supra) animals transgenic for the unc-54/β(1-42) minigene construct contained muscle-specific deposits of anti-β peptide immunoreactivity. Non anti-β-immunoreactivity was observed in control animals (i.e., non-transgenic, wild-type C. elegans tested concurrently). Although C. elegans is reported to contain a homolog of the β-amyloid precursor protein gene (See, Daigle and Li, Proc. Natl. Acad. Sci., 90:2045-2049 [1993]), this sequence does not contain an apparent β-peptide domain. Thus, it would not be expected, nor did it show, cross-reactivity with the anti-β antibody used in these experiments. For dually transgenic animals, immunoreactive deposits accumulated in the body wall of the animals.

Next. to determine whether the immunoreactive deposits observed in the unc-54/ β -(1-42) strains displayed the tinctural properties of classic insoluble β -amyloid, transgenic

- .-... - 961

strains were fixed and stained with thioflavin S, a fluorescent amyloid-specific dye, as known in the art, and described by Guntern and Bouras (R. Guntern and C. Bouras, Experientia 48:-10 [1992]). Thioflavin S-reactive deposits were found in all strains containing the $unc-54/\beta-(1-42)$ minigene constructs, but not in control wild-type animals.

5

Whole mounts of fixed dual transgenic animals stained with coumarin-phallacidin (a muscle-specific probe) and anti- β peptide antibody. Muscle-specific β peptide deposits were observed, and showed a qualitative correlation with the level of GFP expression in these muscle cells (*i.e.*, cells that contained more β -peptide deposits were more green than cells with less β -peptide deposits).

10

In addition, Congo Red and Chrysamine G, two dyes known to interact with β -amyloid and have been reported to interfere with its aggregation are tested on the dually transgenic animals. Anti-oxidants (e.g., vitamin E and ascorbate) are also tested. In these experiments, interference with aggregation of β -peptide, or interaction with β -amyloid are observed.

15

20

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

- 33 -

10

20

CLAIMS

What is claimed is:

5 1. A transparent animal comprising:

- (a) a toxic transgene;
- (b) an inducible promotor operably linked to a reporter gene, wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable *in vivo*.

2. The transparent animal of claim 1, wherein the inducible promoter operably linked to the reporter gene is stably integrated into the genome of the animal.

- 3. The transparent animal of claim 2, wherein the inducible promotor is induced by disruption of cellular function.
 - 4. The transparent animal of claim 2, wherein the inducible promoter is a heat shock promoter.
 - 5. The transparent animal of claim 2, wherein the toxic transgene is β -peptide.
 - 6. The transparent animal of claim 2, wherein said transparent animal is selected from the class Nematoda.
- 7. The transparent animal of Claim 6, wherein said transparent animal is Cuenorhabditis elegans.

8. The transparent animal of Claim 2, wherein said genome comprises SEQ ID NO:5 and SEQ ID NO:8.

- The transparent animal of Claim 2, wherein said reporter gene is green fluorescence protein (GFP).
 - 10. The transparent animal of claim 1, further comprising a heterologous gene marker.
 - 11. A method for producing a dually transgenic non-human animal comprising:
 - a) providing:

10

15

20

25

- i) a first and second non-human animal; and
- ii) a first toxic transgene; and
- iii) a second transgene, comprising a reporter;
- b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal, and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and
- c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic animal, wherein said toxic transgene and said reporter are expressed.
- 12. A dually transgenic non-human animal produced according to the method of claim 11, wherein the toxic transgene is β -peptide.
- 13. A dually transgenic non-human animal produced according to the method of claim 11, wherein said non-human animal is transparent:

WO 98/28971 PCT/US97/23819

5

10

15

20

25

4.

The dually transgenic non-human animal of claim 13, wherein said non-14. human animal is a nematode. The dually transgenic non-human animal of claim 14, wherein said 15. nematode is Caenorhabditis elegans The method of claim 11, wherein said reporter is green fluorescent protein. 16. The dually transgenic non-human animal produced according to the method 17. of claim 11, wherein said first transgene comprises pCL25. The dually transgenic non-human animal produced according to the method 18. of claim 11. wherein said second transgene comprises pCL12. The dually transgenic non-human animal produced according to the method 19. of claim 11, further comprising a heterologous marker gene. A method for testing compounds for toxicity, comprising: 20. a) providing: i) a dually transgenic non-human animal expressing a toxic transgene and reporter gene operably linked to a promoter inducible by the toxic transgene; ii) a composition comprising a test compound in a form suitable for administration such that said compound is bioavailable in the cells of said non-

b) administering said test compound to said non-human animal.

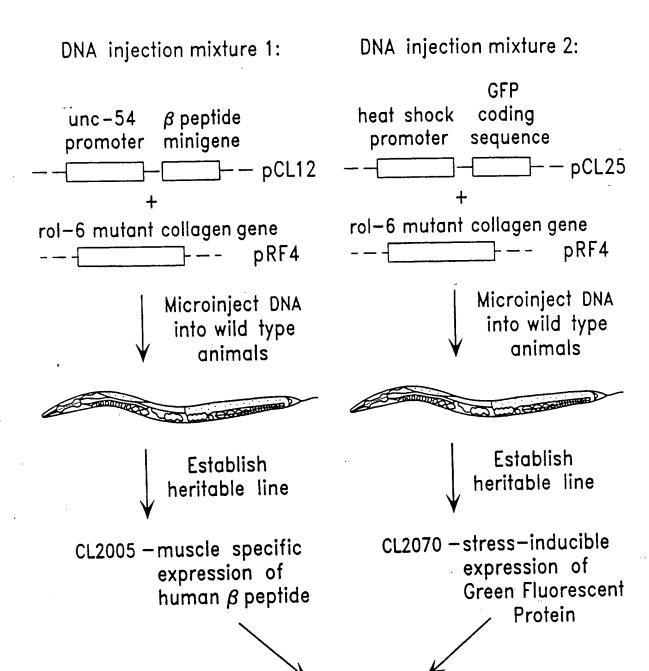
human animal; and

WO 98/28971 PCT/US97/23819

21. The method according to claim 20, wherein the toxic transgene is β -peptide and the reporter gene is GFP.

5

- The method of Claim 21, further comprising c) measuring a reduction in the fluorescence of said non-human animal and thereby identifying a compound as therapeutic.
- The method of Claim 20, wherein said compounds inactivate said β-peptide
 expressed by said dually transgenic animal.



Mate lines together, recover dual transgenic animals by classic genetic techniques

Fig. 1

Fig. 2A
Fig. 2B
Fig. 2C
Fig. 2D
Fig. 2E
Fig. 2F
Fig. 2G
Fig. 2H
Fig. 2I
Fig. 2J
Fig. 2K
Fig. 2L
Fig. 2M
Fig. 2N
Fig. 20
Fig. 2P
Fig. 2Q
Fig. 2R
Fig. 2S
Fig. 2T

> Fig. 2

<u>-</u>

·	3/	⁷ 68	-2A
11 80	I 160	240	11 1 320 320 Fig.
Alu I Mbo III Mbo III Mbo III Mbo III Mbo III Mbo III Hind III Mbo III Mbo III Hind III Mbo II	4 25 ATCCCATTC TAGGGTAAG	Alwn I N I TCAGATACTGCCA AGTCTATGACGGT/ AGTCTATGACGGT/	Mbo II Mbo II Mbo II Mbo II Bg! Bg! AGAAAAACGTTTGAAGAAAGAAGATGAAGAAGATTTTTTTT

W 0 33.255		4/68		2B
320	e I AAAC 400 TTTG 5 5 MSe I	pg I inc II ittaa 480 ittaa 480 itaatt itaatta itaatta	I Mbo II GCTG 560 CGAC 560	Fig.
315 310 310 311 311 311	M	H H H TTTCCCGTTC	Mn I I I I CAAGAGGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
299 308	L IGCAACTGA ACGTTGACT	AGAGTAATGC	c I Mbo II GTGGAGATTA CACCTTCTAAT 539	
5 - 291	Mse I Bsm CTTAAAGGGAA GAATTTCCCTT/ 368 375	CTCTGTAAAA GAGACATTTT	I Sec Hph I GTGGTGAACCG CACCACTTGGC	
5 288	МЬО II В 5 + В I 1 - 1 ТТССААСАА АААССТТСТТ 358 359	TTCACTGGTT AAGTGACCAA	MGB Squ3A I Mbo I Bst7 I Bgl II Mbo II 	
27.5	I GGGGTCTATA CCCCAGATAT	TTTCGGTTT AAAAGCCAAA	Mse I GCTTAAAATC CGAATTTTAG/	000
261	MDe II. AGAAGTAACAO TCTTCATTGTO	CGTCACTCATATA	TTGAATTI FAACTTAAA	
 -	GCACAAGTTA CGTGTTCAAT	Hinf I Hinf I AGGGATTGAATCCT TTCCCTAACTTAGGA		
243	ATTCATT TAAGTAA	AAAGGGA TTCCCT	a I IAGG ATCC	481

	5,	/68		2C
1	640	720	800	Fig.
519 519 520 520 520 520 520 520	Sau3a I Mbo I Dpn I BsiY I BsiY I Bgl II Mse I Ase I Bgl II Mse I Ase I AGATCTGTITAGAACCATTAATCAAACTGGTATTTTCACTGGTTGTAAACATTTTTTTT	Sausa I Mbo I Mbo I Mbo I Hinc II Rsa I Ecor I Opn I Opn I ATCACTGTTCTGCACTTTTTTCCGATCTTACCAACTCC ATCACTGTTCATTTTTTTTTT	SCR II HECOR II HECOR II BEN I BEN I BEN I BEN I MBO II CTATCTCTGGTTTTTTTTTCTTCGAAAAGAAGCAAAAGAAGCAAAAAAAA	22

6/68	
096	Fig.
Scrf I No I I	CCGAACGCCACGGTTGTAGGAGGACGGCTCCTTCGTCCCT 925 933 945 925 936 945
722 ACT ACT A T T G A T U A I U	∀⊢− ∞ '

```
7/68
                                   Bsp
                                                           1109
                                             N a
Ban
                  997
997
998
                            998
998
Hae
Hae
Eae
                   0000
0000
000
                 981
981
982
   Sec I
                                                                                        O
                                                                                     EcoR
        Sfan
                                                                                        Stan
```

WO 98/28971	8/68		<u> </u>
	0/00	_	୍ବ
1200	I 1280	1360	Fig.
	1 I I I I I I I I I I I I I I I I I I I	1 88 0 000 • • • • • • • • • • • • • • • • •	
GGTT CCAA	Sau3A Mbo 1 Dpn 1 CAGAT GTCTA 1278 1278	i A P 12 P 12 CGA CGA 54	→
ACACACACACACACACACACACACACACACACACACAC	TCATC AGTAC	A PO	
S S S	A A A C	TCC AGG	 !
CAGA GTCT 111	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	12 I I AAA AAA	N N N N N N N N N N N N N N N N N N N
AAAC		11] 524 524 6CT(CGA(
AAAG.	S f c S f c	SPH SPH SPH CAT GTA 337	Alu I
66 T C A C C	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	CO CO FA	V -
- A A T T A A A T T	ATC - TA - CO - C	Fok 1 CATC GTAG 1330	
- AH-	A666 T666	II TCAA AGTT	İ
CATC	AGTA	Mbo CTCT GAGA/ 132	. i
AGTA	AACA	ATT •	p // E⊢-
AGT	GCGGTG	₹⊢	₩
ATGA	ဖွဲ့ပ	4 - 5	Mpo
- 66 A T	11 6651 .cc.A	D D D D D D D D D D D D D D D D D D D	; }
11CA AGT 33	N I I I I I I I I I I I I I I I I I I I	GACT GACT	
ACT	P H in H i	I ICT VGA 99	
- ACA - 1- 1- 1- 1- 1- 1- 1- 1- 1- 1- 1- 1- 1-	A + + + + + + + + + + + + + + + + + + +	Mn I FGCCT ACGC 129	
+ i de TCCG AGGC	TAATC.		\ \
Pep- GAAT CTTA	00 A	11.29 12.99 12.99	
1	II AAGG1 TTCC,	10100 66AGC 1	
CGAT CCTA 112	h111 AACA TTGT 01	SO SO STATE	
Stal ATGC	1 th CAA GTT 120	90	1

		9/68			5 <u>G</u>
1440	1520	1600 le II	A 1680	1760	Fig. 2G
CACCCCTATTTTGTTATTATCAAAAACTTCTTTATTTCTTTGTTTTTTTT	MSe I Ase I	TCTGT AGACA	GTTACGTTTCAGTTTATGACCGCAATTTTTATTTCTTCGC GAATGCAAAGTCAAATACTGGCGTTAAAAATAAAGAAGCG GAATGCAAAGTCAAATACTGGCGTTAAAAATAAAGAAGCG 1643	a III TGCTCATCGT ACGAGTAGCA	1697 1707

289/1		10/68			2H
	1840	1920	2000	2080	Fig.
· I IUM	THITGGGGGTITCCCCTATIGITTGTCAAGAGTTTCGAGGACGCGTTTGTTGTTAGAGGGGGGGG	Mbo II Nbo I Mbo I Nbo I Dpn I CAGGAAGAGGTTTGGGTTTGAGGCTCAGTGGAA GTTCTTTCTAGCCTTCTTCCAAACCCAAACTCCGAGTCTT 1885 1907	1883 1890 h	1975 Mse I Dra I 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Msp. I Mse. I
1687 1698 1689 1698	894 1 1 AATTTTCCTGCTTTTGCTI TTAAAAGGACGAAAACGAA	1772 Sfan I Hgia I Bsp1286 I Bsp1286 I AAAGAACGATTTTAGTGAGCACGATGCAAGAAA AAAGAACGATTTTAGTGATTCATGAGAAA	18 SGAGTA	ACICAICIICAACIA	04 2016 HinP I BstU I

-- (0111 F 36)

		11/68		21
2160	2240	2320	2400	Fig. 2
FCOR I Hha I Ha I Ha I Ha I Ha I Ha I Ha I H	2091 Tab Hinf Hinf Hinf Hinf Hinf AAATACTAAG	179 TTTCAAATAAAAACAAAGAAA	Mge III Age I Hinp I Hoa I Hha I Hoa I Hha I Hoa I Hha I Hoa I Age III Hoa I Hha I Hoa III Hoa I Hoa III Hoa I Hoa III Hoa I Age III Hoa I Hoa H	534

...

	12/68	23
AAAGCCAGAGACGCTACTCCAC 2480	Hph I Mn! I NSPH I I NSP7524 I I I NSP7524 I I I NSP7524 I I I I NSP7524 I I I I I I I I I I I I I I I I I I I	Hga I Bstu I Hinp I Hha I
GIGGACTGTATTCTGTCATTTAGGGCA	Squye I Bstu I Hink I Hink I Hhh I Hh I Hhh I Hhh I Hhh I Hhh I Hhh I Hh I H I H	o
I ATAGTTATA TATCAATAT	Sfill RSall Sall Spill RSall Spill RSall Spill RSall R	MSP I SCRF I NC! I BCN I BCN I HPG II MGB III

	13/68	3	2K
2640	2720	2800	Fig.
GTTCGGGCAGTCAG	Rsq I	MSe I Afl II Sfan I Sfan I Hae III FRU4H I ECO109 I FRU4H I ECO109 I FRU4H I ECO109 I Hae III ACCGCACAGATGCGTAAATACCGCATCAGGCGCTTAAGGCCTCGTGATACGCCTATTTTTATAGGTTAAT ACCGCACAGATGCGTAATATTTTATAGGTTAAAATATCCAATTAGGTTAATAGGTTAAAAATATAGGTTAATAGGTTAATAGGTTAATAGGTTAAAAATATAGGTTAAAAATATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAATAGGTTAAAAAA	2763 NIG IV BStU I

	1	4/68		ت
			20	g. 2]
2880	2960	3040	34 I I I I I I I I I I I I I I I I I I I	Fig
Hha I Last CCACCGTGAAACAAAFAAAA AAAAA AAAAA AAAAA AAAAAAAAA	SSP L SSP L AAATGCTTCAATAATATTGA ATTTACGAAGTTATATAACT	TATTCCCTTTTTGCGATTTTGCCTTCCTGTTTTGCTC ATAAGGGAAAAAGGCCGTAAAACGGAAGGACAAAAACGAG	Sau3A I HgiA I Taq I BStY I Mbo I Mbo I Mbo I BSP1286 I Mae III A W I NSPB II A W I	Wse I Hinr I

	15/68		2M
3200	3280	3360	Fig.
CCTTGAGAGITTTCGCCCCGAAGAAAGGTTACTAGTGATGAGACTTTTTTCAAGATTTTCGCCGGTATTATCCCGGAACTTTTCGCCGGCGCGCGTATTATCCCGGCAACTTTTCAAGAGTTTTCAAGACGTTATTATCCCGGCGCGCGC	Fnu4H I Dde I Sca I No Sca I N	T T ACG	

WO 98/28971			ワ
	16/68	3	$\frac{1}{2}$
SA I I I 3440	I 3520	3600	Fig.
CTTACTTCTGACAACGAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTG 3 3415 3422 3376 3383 3393 3440	7 00 4	MSP Hpa Scrf Nci I Mae I Bcn I Mae I Bcn I HGATTAATTGACGCTTGATGATGAGTCGAAGGCCGTTGATGATGAGTCGAAGGCCGTTGATGATGAGTCGAAGGCCGTTGATGATGAGTCGAAGGCCGTTGATGATGAGTCGAAGGCCGTTGATGATGAGTCGAAGGCCGTTGATGATGATTGAT	35 35 35 35 35 35 35 35 35

	17/6	8	20
10 I 3680 80	3760	3840	다. 11년 11년 11년
GTCTC CAGAG 36	CAACT GTTGA	TACTC ATGAG	II DIN
600166 600A000	e I nf I GTCAGGC CAGTCCG 48	CCAAGII GGIICAA	
Pd II	CGGGGA Hill CCCCT 37 37	TCAGA	1
ACCTC 3659	ACACGA TGTGCT	IGE III TAACTG	Squ3A Mbo I Dpn I Alw I BstY I
 TAAATC ATTTAG	TTATCT.	Mae CATTGGTA GTAACCAT	MbG
TGCTGA ACGACT	AGCATC.	Mse I GATTAAG CTAATIC 3807	Mae Sau3A Mbo I Dpn I Bs+Y I
GGTTTAT	I CCCCGTA GGGCAT	I I TCACT AGTGA	Mse I
I TGGCT ACCGA	Mn I 	Mn IV	_ <u> </u>
TCCGGC AGGCCG AGGCCG 3629	I C I	e I GAGAT CTCTA	\ \
Hae II Bg! I CCCCCT CCCCCT 3623 3623	Hae II Sau96 I I a IV I C C C C C C C C C C C C C C C C C C C	0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 0 0 0	Se I
HinP I GCCCTC GCCCTC 3616 3616	6CACT 60 2 31	TAGACA ATCTGT	3
I I ACTTCT 1GAAGA	ATTGCAC TAACGTC 3692	ACGAAA.	
Sau96 Ava I Ava I CAGACC STCTGG 3605 3605	GGTATC/ CCATAG1	Fok I GATGA CCTACT 3763	
ACG Frs 1	00 00	A T A	1

WO 98/28971		
	18/68	2 12
3920	4000	I 4080
Hph I I I I I I I I I I I I I I I I I I I	886 3898 3898 3898 Sau3A I Mbo I Mbo I Mbo I Nbo I	BStU I Fnu4H I Thill II hha I Hha I Bby I Tthill II hha I Hha I Bby I Tthill II Hha I Hha

-- (-) (1 - 76)

##
Hinp I I I Hinp I I Hinp I I Hinp I I Hinp I I Hinp I I Hinp I I I I I I I I I I I I I I I I I I I
Hinp I Hi
Hinp I I I Hinp I I Hinp I I Hinp I I I Hinp I I I Hinp I I Hinp I I Hinp I I Hinp I I I Hinp I I Hinp I I Hinp I I Hinp I I I Hinp I I I Hinp I I I Hinp I I Hinp I I Hinp I I I I I Hinp I I I I I I I I I I I I I I I I I I I
HinP I Hha I Hha I Hha I Hha I Hha I A 118 A 118 A 118 Mul I Mul I Mul I Hune II I Hune III Hune II I Hune III Hune III Hune III CAAGAGGGATACCAATACTGTTACGAATACTGTTACGATGGTTACGT
Hinp I Hinp I Hinp I Hinp I Hinp I 4118 4118 4118 4118 4118 4118 Who I I Who III Hinp I I I Hor I I I Hor I I Hor I I I I Hor I I I I Hor I I I I I I I I I I I I I I I I I I I
HinP I A 118 A 1198 A 198 A 2 198 A
Hinp I Hha I Hha I Hha I His 4118 4118 4118 4118 4118 4118 4198 1 Hha I Hall I
TTCAGCAGA AGTCGTCT GCCTACATA CGGATGTATA CGGATGTATA CAAGACGAT
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
4020 4022 4022 614401 61460 4100 4100 7111111111111111111111111111
CCTAGACCTTGACTTTACCAACCTTGACTTTACCAACCTTTACAACCTTTACCAACCTTTACAACCTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTACAACCTTAC
CTTTT GAAAAA GCACACACACACACACACACACACACACACAC
ACCAACT TGGTTGA GCCACCA CGGTGGT CTATTCA

	20/68		2R
4400	4480	4560	I I Fig.
1 I I I I I I I I I I I I I I I I I I I	T T T O	I CTCGTCAGG GAGCAGTCC	NSPH I NSP 7524
AGCATTGAGAAAG TCGTAACTCTTTC 43	Mnl I HinP I Alu GCGCACGAGGAGC CGCGTGCTCCCTCG 1459 447	I SfdN ATTTTGTGATG TAAAACACTAC 4548	Hae III Hae I
ATACCTACAGCGTG TATGGATGTCGCAC	TCGGAACAGGAGA AGCCTTGTCCTCT	Taq I Hga I CTGACTTGAGCGTCG GACTGGAACTCGCAGC	Hde III Scrf I Ecor II Bstn I
Dde I Dde I TACACCGAACTGAG ATGTGGCTTGACTC	p I Fnu4H I d II GGTAAGCGGCAGGG CCATTCGCCGTCCC 30 4437	Mn I GGTTTCGCCACCT CCAAAGCGGTGGA 452	Hae III Fnu4H I BstU I
I TTGGAGCGAACGACC AACCTCGCTTGCTGG	Ms HP GGCGGACAGGTATCC CCGCCTGTCCATAGG 44	TTTATAGTCCTGTCG	
HgiA I Bsp1286 I ApdL I GTGCACAGCCCAGCI CACGTGTGTCGGGTCG/ H 321 4331	TTCCCGAAGGGAGAAA	SCFF I ECOR II BS+N I CGAAACGCCTGGTATC CCTTTGCGGACCATAG 4488	8

```
21/68
                                                                                                                                                                  Fig.
                                                                                                       0000
0000
0000
                                             TCGCCGCAGCCGA
AGCGGCGTCGGCT
                                                                                               Hae
                                                                                                                                      47964796
44
603
4635
635
635
635
635
635
                                                                        Gd:
Eae
                                                                                 BstC
HinP
Hha J
                                                                                                                      4781
      4609
4609
4609
4609
                                                                                                                            61
762
762
                                                                                                  MulI
                                         Hinf
                                                                                                                                           4736
     4566
```

	22/68	2T
4880	4960	म
AA-	12 11 13 13 13 13 13 13 13 13 13 13 13 13	∞ ∞ ∡

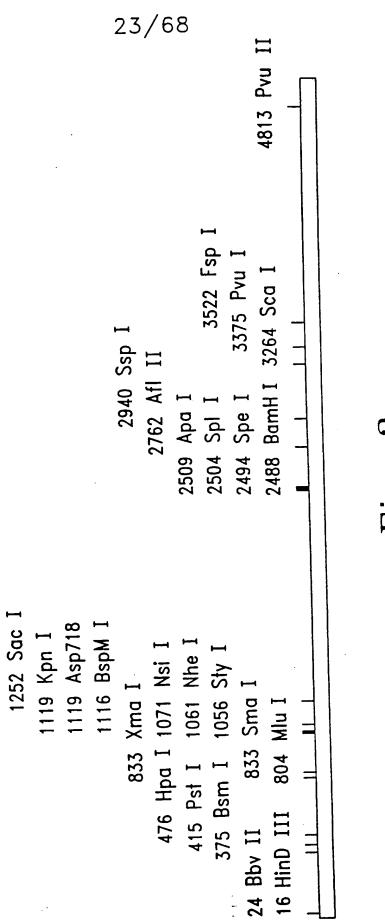


Fig. 3

909

<u>a</u> | a ACC GAT dsb a l a phe thr GCA GGT gly phe 6TGa l a lys leu val CAA AAA TTG pro CTG GCA CCA <u>a | a</u> leu , gln 51 31 phe 1101 / ATC TTT his 1161 / CAT CAT his glu val phe GAA GTT bhe . 110 l eu CTG TCA GGA TAT tyr CCA a l a gly CTG leu ser GAC 116 l eu asb CAT 1 AAG GTT / d / his 1 y s 2 1 CGA arg 41 glu phe 1191 / ATG CAT met his GAA TTC

GTC ۱۵۸ gly val GTG GGC GGT GTT gly leu met val gly ATG CIC GGA TCA AAC AAA GGT GCA ATC ATT <u>.</u> <u>e</u> <u>a l</u>a gly lys dsn ser 199 GAT GTG

gly 61 asp val 1251 /

Fig. 5A
Fig. 5B
Fig. 5C
Fig. 5D
Fig. 5E
Fig. 5F
Fig. 5G
Fig. 5H
Fig. 5I
Fig. 5J
Fig. 5K
Fig. 5L
Fig. 5M
Fig. 5N
Fig. 50
Fig. 5P
Fig. 5Q
Fig. 5R

> Fig. 5

26/68

80 22 22

```
Fig. 5B
                                              27/68
                                                         160
                                       RSa
NIa I
KPN
Ban
ASP Z
                                                                                                       98
                                                                     ထထထထ
     0
                                                                                           Hph
32
                                         Sau96
NIA IV
Ava II
PpuM I
Eco0109
                                                                    ageatageatagageattt
ttgtattgtatettgtagaga
```

	28,	/68		2C
 	320	400	I 480	Fig.
	HGCACATACGGAAAACTTACCTTAAATTTATTTGCACTACTGGAAACTACCTGTTCCATGGCCAACATTTGTTGCATACTACTTTTGATGGACTACTGTTTTGATGGACTACTGTTTTGATGGACTACTTTTGATGGACTACTGTTTTTTTT	III ACTTTT TGAAAA	RSO I GAAGGTTATGTACAGGAAAGAACTATTTTTCAAAGATGACGGGAACTACAAGACGTGCTGAAGTCAAGTTTGAAGG CTTCCAATACATGTCTTCTTGATATAAAAAGTTTCTACTGCCCTTGATGTTCTGTGTGCACGACTTCAGTTCAAACTTCCAAACTTCAAAAAAAA	014

		29/	⁷ 68		 5D
!	560	→ Thr 167 AC 640 TG 10	720		Fig.
	bo II AAGAIGGAAAC TICTACCITIG	XCQ I HPQ I I HPQ I HPQ I HPQ I HPQ I HPQ I HPQ I I HPQ I I HPQ I I HPQ I HPQ I HPQ I I HP	Mae I CAACTAGCAGACCATTATCAACAAATACTCCAATTGGCG GTTGATCGTCTGGTAATAGTTGTTTTATGAGGTTAACCGC	Sau3A I Mbo I Dpn I Alw I Bsty I Aya II Bsty I Bsty I Bsty I Bsty I	AGACAACCATTACCTGTCCACAATCTGCCTTTCGAAAGATCCCAACGAAAAGAGACCACATGGTCCTTCTTGAGT TCTGTTGGTAATGGACAGGTGTTAGACGGGAAAGCTTTCTAGGGTTGCTTTTCTCTCTC

	30/68		5E
 	880	096	Fig.
754 755 760 761 761 761 761 761	Fourth I	NSPB II FRU4H I Squ96 I Hge III Hge III Hge III Hge III Hhg I Hg I Hhg I Hhg I Hg I H	NIGITION.

	31	/68		- 5F
1040	se I 1120 120	1200	1280	Fig.
524 I GCTCT CGAGA	Hph I Mde III Xmd I I I I I I I I I I I I I I I I I I I	Mni HgiA I Bsp1286 I ATAAAAATTGTGCTCCCT TATTTTCAGCTTTTTTTAACACGAGGA 1139 1151	RSO I TTCTGTGTACACTTCTTATGTTTTTTACTTCTGATAATTTTTTTGAAACATCATAGTATTTTTGGAAAAAACGCGCACAAAATA AAGACACATGTGAAGAATACAAAAAAATGAAGACTATTTAAAAAAACTTTGTAGTATCTTTTTTGGCGTGTGTTTTAT AAGACACATGTGAAGAATACAAAAAAAAATGAAGACTATTTAAAAAAAA	, 🛎

W 0 30/203 -			
	32/68	3	<u>ت</u>
1_	-		12
360	1440 i A I p 1286 1520	1600	
259		152 152 •	Gd : Faga Fa
- 0.0	GAT	AAAC	1
A H	SAC SAT TAAT	1 GA ACT	
100 000 100 000	A A A G C A A C C A C A C C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C	T T A A T T	1
7 TGAC ACTGC 134 134 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ATT FAA. 5 ACA TGT	A T A	1
134 134 134 138	A A A A A A A A A A A A A A A A A A A	AAC.	
GAG/ GAG/ 9341	AAA TTT	SAAG	į
- 3 - 8 - 6 - 6 - 6 - 6 - 6 - 6 - 6 - 6 - 6	ATG. TAC.	TAC.	. !
6A50-1-	AAA CTT GAA	h I TGAG' ACTC/	
332	TTT TTT AAA	HPF 	1
- 415	GCAACT GCATT GCATT	166A/	1
11CG AAGC 25	CAAG GTTC	I I I I I I I I I I I I I I I I I I I	
13 13 13 13 13 13 13 13 13 13 13 13 13 1	TAT TAT	31 Dde CTC GAG	i • I
ATA	A P D D D D D D D D D D D D D D D D D D	(n() m	
1111 AAA.	AAAA.	T T A A	j
TAAT	SAG CTC	GGGT) 1
00 00 00	TG VAC	A A	
ACT CTO O	AA 1	့ ဗွ်ပ္ပ	1
A T A	ATA H 1	1464 o II AGAA	2
AAT AAT		Mb - CTT-	154
TTCAG AAGTCG	La. 00	A MO UCITI	ו נאוני
l ⊢ ≪ Ω	<u> </u>	Y AL	<u>c.t.</u>
17 A T G C G C C C C C C C C C C C C C C C C	AAAA TTTT	<i>-</i>	
ACA 12	A0. F0.	I I	
	55	S - A - N	
ATATC- - TATC- 	ST ST	∢	!
CC11	16C1	ဗိုင်	I

```
33/68
                                                                              1760
          1680
                            680
680
680
                                                                                         1758
- Mde
Spe 1
                     1674
                                                                  NSPH I
                                                                                           1730
                                                                                                                                         Hga
Bs†U
                                                                          Hbh
                         637
       Xmn
                                                                           Hbh
                                                                                            703
1704
704
1705
                                                                Bstu
Hinp
Hha
Bstu
                                                                                                                                 Scr
Moori
Boor
                                            Squ96 1
Squ96 1
N|a IV
Eco0109
Bsp1286
                                                                                                                1689
1689
1689
1689
                                                                                               1690
689
689
                                                                             Hae
                                                                     Ban
                                                                                              1681
1681
1682
                                                                          Sf.
Hde
```

	34/68		19
1840	1920	de I 2000 000	Fig.
ACAGCTTGTCTGTAAGCGGATGCCGGAGCAAGCCCGTCAGGGCGCTCAGGGGGTGTTGGCGGGTGTCGGGGGTGTGGCGGGTGTCGGGGGTGTTGGCGGGTGTCGGGGGTGTGGCGGGTGTCGGGGGTGTTGGCGGGTGTCGGGGCTGTTGGCGGC	783 783 161/	Mni I Squ96 I Eco0109 I III Hae III I Hae III FAGGGCCTCGTGATACGCTTTTTTTTTTTTTACGTTTACAGTACTATTACCAAAGTTTCCCAATTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTTTACAGTACTATTACCAAAGTATTACCAAAGTATTACAGTACTATTACCAAAGTATTACCAAAGTATTACCAAAGTATTACCAAAGTATTACCAAAGTATTACAGTACTATTACCAAAGTATTACCAAAGTATTACAGTACTATTACAAAGTATTACCAAAAGTATTACAATTACCAAAAGTATTACAAAATATACCAAAAGTATTACAAATATACCAAAAGTATTACAAATATACCAAAAGTATTACAAATATAAAATATACCAAAAATATACCAAAAATATACAAATATACCAAAAATATACCAAAAATATACAAAAATATACAAATATAAAAATATACAAAAAA	1942

		35	/68			51
	ATGT 2080	1CCGT 2160	Sfan I AAGATGC 2240 FTTCTACG	II O9W	TCGCCCC 2320 AAGCGGGGC 2320	PIT PIT
.	TCTAAATACATTCAAATA ,agatttatgtaagtttat	ATTCAACATTT TAAGTTGTAAA	I GAAAGTAAAA ACTTTÇATTTT	23	TGAGAGTTTTC ACTCTCAAAAG	
	TTT	Mbo II AGGAAGAGTATGAGTATTCAACATT	33 Hph GAACGCTGGT CTTTGCGACCA	2223 Squ3A I Mbo I Dpn I BstY I	AIW I ATCCT TAGGA 12298	2298 2298 2298
	TTGTTT AAACAAA	Mbo TATTGAAAAAGGAA	HP 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Sau3A I Mbo I Don I	Y I NSPB II I I I I I I I I I I I I I I I I I	2880
	GCGCGGAACCCCTA CGCGCCTTGGGGAT	SS	212	AAOAOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOO	Taq I BS ¹ Y I 10 III A W I 11 ACATCGAACTGGATCT 11 ACATCGACTGGATCT 11 ACATCGACTTGACTAGA	122
VI DIN	BstU 1 linp I tha I GGGGAATGT CCCCTTTACA	2035 2035 2045 2041A	STATTAC STATTAC 4H I	TGCGCGTAAAAC ACGCCGTAAAAC 2182	6 1 MC	2
1943	rggcactttt accgtgaaa	_ C	AGACAA LAACTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOT	TTCCCTTTTT AAGGGAAAAA	Hgia Bsp128 Bsp128 ApdL CTTGGTGCA	2255 2255 2255
	MGB II Aha II Aat II TAGACGTCAGG	2004 2004 2005 2005 N BS	ATCCGCTCATG TAGGCGAGTAC 1 2087	GTCGCCCTTA CAGCGGGAAT	Sau3A Mbo J Dpn 11	444

36/	⁷ 68		2K
2400	2480 I	2560	Fig.
BS + U I	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	H I PVU I CCAACTTACTTCTGACAACGATCGGTTGGTTGCTAGGTTGCTAGGTTGCTAGGTGTGTTGCTAGGTGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGCTAGGTGTTGTTGCTAGGTGTTGTTGCTAGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGT	

,	37/68		2T
2640	2720	2800	Fig.
Sau36 I Sau36 I Alw I Alw I Alw I Alw I Alw I Alw I Alu I	GAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTC	MSP II Scrf II Nci I MSe I Fok I Ava II Hha I Mae I Pok I Ava II Hha I Mae I Pok I Ava II Hha I Mae I Pok I Ava II Hha I CGAACTACTTACTCTAGCTTCCCGCCAACAATTAATAGACTGGAGCCGCTATTTCAACGTCCTGGGAGCCGA GCTTGATGAATGAGGCCGTTGTTAATTATCTGACCTCCCCCTATTTCAACGTCCTGGTGAAGACGCCGA	2

	38/68		5M
	45 Pie I Hinf I GGAGTCAGGCAACTATGGATGAATAGACA 2960 CCTCAGTCCGTTGATACCTACTTGCTTTATCTGT CCTCAGTCCGTTGATACCTACTTGCTTTATCTGT 2928	MS8 I Dra I I I I I 3040 TCTGGTTCAAATGAGTATATGAATCTAACTAA 3039	Fig. 5
SP I MSP HPD CGCTGGTTTATTGCTGATAAATCTGGGCCGGCCGGCCGACCGA	809 809 Wn! I GGTAAGCCTCCGATAGCATAGATAGATGTGCTGCC CCCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCC	I Mn! I NIG IV WSB I MGB III BGN I AGATAGGIGCCTCACTGATTAGCATTGGTAACTGTCA TCTATTCGTAACTGACTGACTGACTGACTGACTGACTGAC	
B91 I HQ6 III HQ6 GCCCTTC GCCCCTTC GCCCCTTC GCCCTTC GCCCTTC GCCCTTC GCCCTTC GCCCCTTC GCCCTTC GCCCCTTC GCCCCTTC GCCCCTTC GCCCTTC GCCCCTTC GCCCCTTC GCCCCTTC GCCCCTTC GCCCCTTC GCCCTTC GCCCCTTC GCCCCTTC GCCCTTC GCCCCTTC GCCCCTTC GCCCCTTC GCCCTTC GCCCTTC GCCCCTTC GCCCTTC GCCTTC GC	III I I TCT	O DOL HA 96	න

•	39/68	5N
11 16AG 3120 ACTC 4	Fnu4H I Bbv I TCTG 3200 AGAC AGAC 3200 3200	Mae III
MS9 I HS9 I ATCCTTAACG TAGGGAATIGC 3111	BS+U I HinP I Hhd I -1 CTGCGCGTAA AGACGCGCATT 3189 3189 3189 3190	ACTCTTTTCC(
NIG III SPH I CATGACCAAA CATGACTTT GTACTGGTTT 095	Squ3A I Mbo I Dpn I Alw I BstY I ICTAGGAAAAA ICTAGGAAAAA 3175 3175 3175	Dpn I Alw I Sp II Pg II Alu I CGGATCAAGAGCTACCAACTCTTTTTCCGAAGG 3
TTTTGATAATCT AAAACTATTAGA 33	Sau3A I Mbo II BstY I A w I CCTAGAGAAC CCTAGAGAAC II 3163 3164 3164 3164 3164 3164	Alw I MSP I MSP I HPG II I I I IGTITGCCGGATC
BS † Y I BS	Squ3A I Mbo I Mbo I Dpn I TTCTAGTT 3156 3156 3156	3 II 50000000000000000000000000000000000
I BS PY I HP AAAGGATCTAGG TTTCCTAGATCC 3065 3059 3066 3066 3066 3066 3069	GACCCCGTAGAA CTGGGGCATCTT	Ns CACCGCTACCA
MSe I Dra J MSe I H H H H H H H H H H H H H H H H H H	Hga I Dde I TTCGTTCCACTGAGCGTCA(AAGCAAGGTGACTCGCAGT(3132	THHIII II THHIII II CTGCTTGCAAAAAAACCACCGCTACCAGG
TAAAACTTCA ATTTTGAAGT	TTTTCGTTCC AAAAGCAAGG	T+h111 1 CTGCTTGCA

		40/68	50))
3280	3360	1 3440	3520 Fig.	Q
32	SAAGAACTCT STICTIGAGA	ATACCGC ATGCCGC ATGCCCC ATGCCCCC ATGCCCCC ATGCCCCC ATGCCCCC ATGCCCC	GCTTGG CGAACC	
AAAAAGG	□ ▼·	TCT AGA	HgiA I Bsp1286 I ApdL I CGTGCACACACCAGCTTGG AGCACGTGTGGGTCGAACC	
GTTGAG	CCCAC GGTG	TAAGTC TATTCAG	HgiA I Bsp1286 ApdL I CGTGCACAC	
3257	HGE I HGE I HGE I CTTAGGCCA ICAATCCGGT 3339	61 60 60 60 60 60 60	- ∀	
CTAGITC 7 7 249 3250 3250	GCCGTAC	Fnu4H I Fnu4H I N I Bbv I Bbv I I CCTCTGCCAGTGGCGATAAGTCGTG ACCGACGACGTCACCGCTATTCAGCACG ACGACGACGTCACCGCTATTCAGCACG ACGACGACGTCACCGCTATTCAGCACG ACGACGACGTCACCGCTATTCAGCACG ACGACGACGTCACCGCTATTCAGCACGACGACGACGACGACGACGACGACGACGACGACG	7ACGGGG	
CCACCAAACAACGGCCTAGTTCTCGATGGTTGAGAAAAGGCTTCC 3241 3247 3257 3257 3250 3250 3250 3250	de I TAGTGTA ATCACAT 322	Fnu4H I Bbv I Bbv I Bbv I CCTGCTGCCACGACGACGACGACGACGACGACGACGACGACGACGA	CGGCTGAACGGGGGG	
CAAACA 324	A-00-W	Alwn 		
GTCGCCAC 3230	ATACTGTCCT TATGACAGGA		MSP I Fnu4H I Hinh I Hinh I Hog I Hha I Hagirage Coccessor Age Coc	
CGAT GGT 32	TACCAA/ ATGGTT	CTGCTA GACGAT	SP I II CGGATAA GCCTATI	
GACGAACGTTTGTTTTTGGTGGCGATGGTCG 3202 3208	HinP I Hha I Hha I Hha I TTGACCGAGAGCGAGATACCAAAT	GTAGCACCGCCTACATACCTCGCTCTGCTAATCCTCTCCTCGTCGCGCAGCGAGCG	MAG II HP HP HP HP HP HP HP HP HP HP HP HP HP	
111111	CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	ATGTACATAC	GACGAT	
3208	SGCTTCA	ACC 500 500 500 500 500 500 500 500 500 5	Inf I SACTCAA	
6 <u>ACGAA</u> 3202	TAACT(ATTGA(GTAGC CATCG	Ple Hinf CATCGACT	

	4080	43/	/68 80 80	v	g. 5R
3971 3976 3976 3976 3976 ScrF I EcoR II BS+N I	Sec I HinP I Mse I Hha I Ase I	4035 4046	pd II cgctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatga gccgagcatacaacaccttaacactcgcctattgttaaagtgtgtgt) }	Fig.
	ACGACAGGTTTCCCGACTGG	,	GCTTTACACTTTATGCTTC	Id III ATGATTACGCC 4173 TACTAATGCGG	4162

HinD III 7 Sph I 13 Pst I BspM I 15 19 Sal I 25 Xba I 31 BamH I 36 Ava I 36 Sma I 36 Xma I 118 Asp718 118 Kpn I 299 Nco I ,326 581 Xca I Ser 65 → Thr 620 Hpa I 633 Ile167→Thr 850 Bsm I 859 EcoR I presumed gGFP-TT 4173 base pairs unique sites Fsp I 2702 Pvu I 2555 1674 Spe I Sca I 2444 1680 Eag I 1684 Spl I Ssp I 2120 1689 Apa I San II 1689 Fig. 1942 Afl II

44/68

		•
	Fig. 7A	
	Fig. 7B	
	Fig. 7C	
	Fig. 7D	
	Fig. 7E	
	Fig. 7F	
	Fig. 7G	
	Fig. 7H	
	Fig. 7I	
	Fig. 7J	
	Fig. 7K	
	Fig. 7L	
	Fig. 7M	
	Fig. 7N	
	Fig. 70	
•	Fig. 7P	
	Fig. 7Q	
	Fig. 7R	
	Fig. 7S	
	Fig. 7T	

> Fig. 7

```
46/68
        Tag
Bs+B
           III DIN
            Hinf
                                         22
22
DZ
```

	47/68	_7B
320	400	년 원
TGTATCTGGGTCTTTTTTTTTTTTTTTTTTTTTTTTTTT	GAAATAAATACCGAACAACATTTGCTCTAATTGTG CTTTTTATTATGCCTTGTTGTAACGAGATTAACAC I I I I I I	Smg I Smg I Sec I Nci I Bcn I Avg I Sec I Sau3A I

```
48/68
           480
Sau96
NI a IV
                       53
53
                452
522
   Hde
Ede
Bd-
                                                      Rsd
NIG
Kpn
Bgn
Asp7
                                           MUMMMMM
                 Sau9
NIa
Ava
             00.
```

```
49/68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fig.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hde Fede No. 1 No. 2 No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        700
700
700
700
MC
                                                                                                                                                                                                                                                                                                                          2000
2000
2000
2000
2000
2000
2000
                                                                                                                                                                                                                                                                           516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hph I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  592
```

	50/	68		7E
800	880 0	4 096	1040	Fig.
Thres Thres Thres ACTIT CACTITATE CATECTITICA CATACT CATAT CAAACT CATATICA ACTITITICA CATECTACE CATECTACE CATECTACE CATACTATICA ACTICACE CATACTACTACTACTACTACTACTACTACTACTACTACTA	763 Affill Affill Affill GGGAACTACAAGACACGTGCTGAAGTTTGAAG GCCCTTGATGTTCTGTGCACGTTCAAGTTTGAAG BS6 856	GATGGAAACATTCTTGGACACAAATTGGAA CTACCTTTGTAAGAACCTGTGTTTAACCTT	XCQ I ACC I ATGTATAC TACATATG 982 982	

VY O 30/20771	51/68	•	7F
1120	1200	1280	Fig. 7
Mbo II Mae I CAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCTGTCCTTTTAC GTTGTAACTTCTACCTTCGCAAGTTGATCGTCTGGTAATAGTTGTTTATGAGGTTAACGGCGATGGAGAAATG 1105	Sau3A I Mbo I Don I Balw I I Balw I I Avg II Avg II I Avg II Avg II I Avg II I Avg II I Avg II I Avg II Fnu4H I Bbv I Alu I Vu II SpB II IWN I II NIG III II AGCTGCTGGGATTACACATGGCATGGA TCGACGACCTAATGTGTACCGTACC		

		52/68		76
1	1360	1440	1 1520 1520 9	Fig.
1208 1209 1210 1210	2 1 AC TG	ပ္ခ်	Hph I Mae III Mul I Kun Hph I Kun Hph I Kun Hph I Kun Kun Hph I Kun Kun Hinf I Kun Kun Hinf I Kun Kun Hinf I Kun Kun Hinf I Kun	Mse I

TTCGTAATAAAAGTCGAAAAAA	1600	
AATTAAGCATIAIIIICAGCIIIIIIGAGCAGCATAAGCATIAIIIICAGCATIAIIIICAGCATIAIIIICAGCATAGCAT		
RSO I GIICTGIGIACACTICITATGITITITITACTICTGATAAATTITITITIGAAACATCATAGAAAAAAACGGCACACACAAAT CAAGACACATGIGAAGAATACAAAAAAAAGACTATTTAAAAAAAAAA	1680	
Hae III Mae II Mae II Aha III NAC III Mae III Aha III NAC III Mae III Aha III NAC IIII	III DI	
MGE II ACGITICAGITITATGACCGCAATTITITATTICTTCGCACGTCTGGGCCTCTCATGACGTCAAATCAATCGCACGTTTGGGGCCTCTCATGACGTTTAGTTCAGAAGCGTGCAGGAGGGGAGGTACTGCAGATTAGTTTAGTTGCAAAGAAGAAGGAGGGAG	1760	
1740 1746 1746 1746 1746 1746 1746 1747 1746 1747 1748 1747 1748 1748 1748 1748 1748		
TCAAGTGAAAGTTTATGAAA AGTTCACTTTCAAATACTTT	1840	
Mni I Tag I TTGCITTTTGGGGGTTTCCCCCTATTGTTCAGGGTTTCGAGGGGCGGTTTTTCTTGCTAGAATCACAAGTATTGAT	1920 Fig.	~

	54/68		71
2000	2080	111	Fig.
AACGAAAACCCCCAAAGGGATAACAAACATTCTCAAAGGTCTTGCGCCAAAAGAACAACGATTTTAGTGTTCATAACTA 1865 1880 1882 Hgia I Nbo II Nbo II Nbo II Nbo II Hph I GACCACGATGCAAAAGGTTTGGGTTTGGGTTTGAAGTTGAAAGTTGAAAGTTGAAAGTTGAAAGTTGAAAGTTGAAAGTTGAAAGTTGAAAGTTTGAAAGTTGAAAGTTGAAAGTTTGAAAGTTTGAAAGTTGAAAGTTTGAAAGTTTGAAAGTTTGAAAGTTTTGAAAGTTTTGAAAGTTTTGAAAGTTTTTT	Mae I Kmn I Spe I	U96 I 96 I 1V 0109 I 1286 I Hinp I 1 Hinp I 1 Hph I Mn! I NSPH I Hpd II Mae 11 Bstu I 11 Hpa II Mae 1286 I Hpa II Mae	0 2104 2114 2123 2131 2139 2150 2159 0 2105 2105 Fig.
AACGAAAACCCC AACGAAAACCCC Sfan I Bsp1286 I GAGCACGATGCAA CTCGTGCTACGTI	TGGAGTAGTGTC] ACCTCATCAÇAGA	Sausenses Sausen	081

```
55/68
                                                                                              Fig. 7J
                                                                       2320
                                                                   Sfan I
                                                                                      2279
2279
                                                  2090
2090
2090
2090
```

	56/68			7K
2400	2480	2560 I	2640	Fig.
TAATGGTTT	TCAAATATG AGTTTATAC	ACATT TGTAA	TAAAAGATG ATTTTCTAC 2637	
BSPH I III NIG III NIG CATGATAA	TCTAAATACAT AAGATTTATGTA	TATGAGTATTCA. ATACTCATAAGT	CGCTGGTGAAAG GCGACCACTTTC 2624	
MSB GCCTATTTTTATAGGTTAA CGGATAAAAATATCCAATT	IV ACCCCTATTTGTTTATT TGGGGATAAACAAATAA	Mbo II AAAGGAAGAG TTTCCTTCTC 2534	TTTTGCTCACCCAGAAA AAAACGAGTGGGTCTTT 2609	
Sfill Squ96 I Fnu4H I Ec00109 I Fnu4H I Ec00109 I Hquart Accept CccccTTAAccccTCGTGATACCCTTTTTTTTTTTTTTTT	II BstU I II Hinp I H	ACAATAACCCTGATAAATGCTTCAATA TGTTATTGGGACTATTTACGAAGTTAT 2	GCATTTTGCCTTCCTG CGTAAAACGGAAGGAC 3	1

	57/68	17	7
2720	2800	2880 Fig.	
11106000 AAAGCGGGG	7 TO	RSa I Mae III ŞfaN I Fok I Sca I Nae III ŞfaN I Fok I SCACTCGGTCGCCGCATACAGGATGCGATG CTTACGGATG CTTACGGATG CTTGGTTGATGACTCATGGTCACAGAAAGCATGTTACGGATGCCTACTACGGATGCCTACTACGGATGCCTACAGGCGCCAGCGCGCGC	

	58/68		ZM
I 2960	3040	3120	Fig.
TGACAGTAAGAGAATTATGCAGTGCATAACCATGAGTGATTGTGGGTTGAATGAA	Squ36 I Squ36 I Squ37 Squ38 I Squ38 I Squ38 I HPQ I Mbo I Mbo I HPQ I Mbo I Mbo I HPQ I Mbo I HPQ I I HPQ I I HPQ I I HPQ I I I I I HPQ I I I I HPQ I I I I I I HPQ I I I I I HPQ I I I I I I HPQ I I I I I I HPQ I I I I I HPQ I I I I I I HPQ I I I I I I HPQ I I I I I HPQ I I I I I HPQ I I I I I I HPQ I I I I I HPQ I I I I I I HPQ I I I I I HPQ I I I I I I HPQ I I I I I HPQ I I I I I I HPQ I I I I I I I I I I I I I I I I I I I	S 50	

```
59/68
                                                                                                                          Fig.
    3200
               3197
3197
                                                                    3273
3273
                                                           TCTCGCGGT
AGAGCGCCA
                                                                                           Ple I
                  152
3153
Bcn
                                                 Bgl
Sau96
Hae I
```

	61/68		.7P
	3680	3760 1	3840 Fig.
3590 3590 3591	CCAACTCTTTTTCCGAAGGGTTC	GCCACCACTTCAAGAACTC GCCACCACTTCAAGAACTC CGGTGGTGAAGTTCTTGAG CGGTGGTGAAGTTCTTGAG	1
3575 3576 3576 3576 3576	1 A U I A U I A T C A G C T A C T C C C A T C C C C	651 Hqe CCGTAGTTAG GGCATCAATC GGCATCAATC 11 11	CTGCCAGTGGC GACGGTCACCG
3564 3564 3564 3567 3565	265 M M M M M M M M M M M M M M M M M M M	MGB I TTCTAGTGTA AAGATCACAT 3723 AIWN I B	2 - DA
3557 3557 3557 7557	NspB II CAGCGGTGGT GTCGCCACCA	AAATACTGTCC TTTATGACAGG	MDB LAATCCTGTT ATTAGGACAA
3533	T+h111 II Bby I T+h111 II GCTGCTTGCAAACAAAAACCACCGCTAC GCACGACGTTTTTTTGGTGGCGATG 1 3609 3601 3609	HINP I Hha I GTAACTGGCTTCAGCAGAGCGCAGATACC/ CATTGACCGAGTCGTCTCGCGTCTATGG A 3681	Mn1 I HETAGCACCGCCTACATACCTCGCTCTGC ACATCGTGGGGGACGAGGGGGGGGGG

	62	2/68		70
 	3920	4000	4080	Fig.
3797 3807 3810 3810 3818 3838 3838 3838 3838 3838	NSPBII FRU4H I FRU4H I Bbv I Bbv I Bbv I Hinf I Hinf I Hinf I Hing I Hin	GCACTCGTAGCATTCGCCCCCCCCCCCCCCCCCCCCCCC	SP I Fnu4H I Pd II CGGTAAGCGCAGGGTCGGAACAGGAGAGC GCCATTCGCCGTCCCAGCTTCGC	

WO 98/28971			
	63/68		7R
4160	4240	1 4320	Fig.
Mul I Hga I Stan I Hga I Stan I NI a IV ATCAGGACCACCCAAAGCGGTGGACTTTTTTTTTTTTTT	NIGHTI NSPASSA I HOB III NSP7524 I HOB I Aff III Aff III	Fru4H I Bbv I HinP I HinP I Hinf I Hhq I Hinf CACCGACCGACCGACCGACCGACCACCACCACCACCACCA	

```
64/68
                                          SCT
BCCT
BCCT
TAN
                                                    N
Ban
Ban
Gd:
```

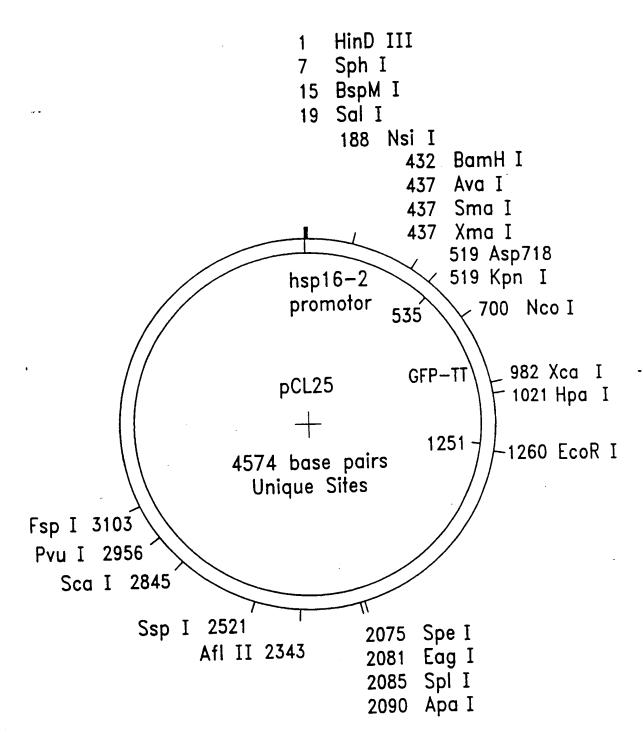


Fig. 8

aggaccagat gggcacgatg gacaagacgg agtcccagga gaccaccagg agaggctggg accaggaagac aggaaaggga ggatgaaaaa aggaatgcgc ggagatgctg 1ggaccacaa accaaatgga accaddacca gaatggagga caactecaaa eccatatgga agatatagaa atticagace agaaggaeea eeaggageee aagatgeega agatgtteaa aaeaeeeeae eaaeaggatg etteaeetgt iggitaaact iggagcagga accgeticca accgigigag acgicaacad cattttcca tgagctggat gccgaaatcg ccaacttcag agtgagtttc agatticcag atgaccetaa etaeggegae gieeggegee tttetaatte actgataatt teaattattt agagteteae tgaagaeatg aaaattitti gtittgtatt tggttfacca tggatgttaa ctttatcatt tcactttttc ttggcactcg tccaaaaatc tcaattaaaa atttctggat acccaggaat gccgaatact gcaaatgccc aggaagagaa aatggagctc caggactcag gatgettate caggacagte atatggagga gctgctggag aggacgggta gatctccagg Highigati tetetititig ecgetgette getttaeagt atcacaatga catecttatg tatatgeatt etttttattg titeetgata titattetet cagatggtat gtccattccc acteatacee cgcaaattcc aattgtagac aaattcatga agagacggaa ggaccaagag ggaaagccag gagaaattcc actgttgaga tggtgttcag ccaccagcac aggaggtgga ttcccagatg aaaactatat aactggaaga acaaggacca gggccgccca acgtccagga agctccacca atatcaatgo aggatccgat attaaaatac aaggagagga cacagggacc gccttcaagg gaggaaatca caagccagcc gacaaccagg aggaaaagat ეეეენნნნნე agagaccatt cacttggacc cagacgicat caaagtteee afggagccac catacticic ggggaaaata atatctggaa gacttettgt atctaactga ctggagccac ggagagcag ccagaaggac aaaaataaaa gacgaagtgc ccacaaggac gacgatggag ggaaatcaac aatattaaac caggacttcc gctggaccag gtggaccac ggaccacgtg ggagctcgtg ggatatggag Hgacggaa ctatatagca atggaggat caagtttcca gaactttett gggttgata atatetttag attgtattt .96 54, 09 301 48 24, œ တ် 7

Fig. 9

DCGPPGAPGSDGKPGSPGGKGDDGERPLGRPGPRGPPGEAGPEGPQGPTGRDAYPGQSG KFPGGGFPDGPFPNGGGPRGGNQCQCTVENSCPPGPAGPEGEEGPDGHDGQDGVPGFD GKDAEDVQNTPPTGCFTCPQGPLGPQGPNGAPGLRGMRGARGQPGRPGRDGNPGMPG **LGAGTASNRVRRQQYGGYGATGVQPPAPTPNPYGGYGASQPAPPEKFPDGIPNGGNQP** PQGEPGLQGYGGAGEDGPEGPPGAPGLPGKDAEYCKCPGREGDAGRSARRHRKFQL MTLTTATSGAIVFSGATLLVSLFAAASLYSQVSNIWNELDAEIANFRSLTEDMWVDMVK

Fig. 10



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶:
C12N 15/00, A01K 67/033, C07K 14/435,
14/47

A3 (11) International Publication Number:

WO 98/28971

(43) International Publication Date:

9 July 1998 (09.07.98)

(21) International Application Number:

PCT/US97/23819

(22) International Filing Date:

31 December 1997 (31.12.97)

(30) Priority Data:

08/778,645 60/080,222 3 January 1997 (03.01.97) US 3 January 1997 (03.01.97) US

(71) Applicant (for all designated States except US): UNIVER-SITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): LINK, Christopher [US/US]: 1837 Spruce Street, Boulder, CO 80302 (US).

(74) Agents: MONROY, Gladys, H. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 29 October 1998 (29.10.98)

(54) Title: TRANSGENIC ANIMALS FOR IN VIVO VISUALIZATION OF BETA-AMYLOID TOXICITY

(57) Abstract

The present invention relates to methods and compositions for visualization of the toxic effects of transgenes in vivo. In particular, the present invention provides methods and compositions for the production and use of transgenic, including dually transgenic, Caenorhabditis elegans for visualization of the toxic effects of \(\beta\)-amyloid accumulation in vivo.

DNA injection mixture 2: DNA injection mixture 1: **GFP** heat shock coding unc-54 \$ peptide promoter sequence promoter minigene - pCL25 - pCL12 rol-6 mutant collagen gene rot-6 mutant collagen gene pRF4 pRF4 Microinject DNA Microinject DNA into wild type into wild type animals animals Establish Establish heritable line heritable line CL2070 - stress-inducible CL2005 - muscle specific expression of expression of Green Fluorescent human & peptide Protein

Mate lines together, recover dual transgenic animals by classic genetic techniques

ir ational Application No PCT/US 97/23819

A. CLASSII	FICATION OF SUBJECT MATTER C12N15/00 A01K67/033 C07K14/	435 C07K14/47	
	•		
According to	International Patent Classification (IPC) or to both national classifie	cation and IPC	
	SEARCHED		
	A01K C07K	ion symbols)	
Documental	tion searched other than minimum documentation to the extent that	such documents are included in the fields sea	rched
.,			
Electronic d	lata base consulted during the international search (name of data b	ase and, where practical, search terms used)	
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the re	elevant passages	Relevant to claim No.
X	CA 2 088 379 A (CANDIDO EDWARD I :STRINGHAM EVE G (CA); JONES DOI 30 July 1994		1-4,6,7. 13-15,20
Y	see the whole document	_	1-23
Y .	LINK, C.D.: "Expression of humbeta-amyloid peptide in transger Caenorhabditis elegans" PROCEEDINGS OF THE NATIONAL ACAI SCIENCES OF USA vol. 92, no. 20, 26 September 1 9368-9372, XP002065639 WASHINGTON US cited in the application	DEMY OF	1-8. 10-15, 18-20. 22.23
X	see the whole document	-/	11-15
X Fur	ther documents are listed in the continuation of box C.	Patent family members are listed	ın annex
- Special c	calegories of cited documents	"T" later document published after the inte or priority date and not in conflict with	ernational filing date in the application but
consi E" earlier filling E' docum which citation O" docum other	nent defining the general state of the art which is not idered to be of particular relevance of document but published on or after the international date ent which may throw doubts on priority claim(s) or his cited to establish the publicationdate of another on or other special reason (as specified) ment referring to an oral disclosure, use, exhibition or rimeans ment published prior to the international filing date but than the priority date claimed.	cried to understand the principle or the invention. "X" document of particular relevance, the cannot be considered novel or cannot involve an inventive step when the different of particular relevance; the cannot be considered to involve an indocument is combined with one or ments, such combination being obvirint the art. "3" document member of the same paten.	claimed invention of the considered to ocument is taken alone claimed invention overtive step when the ore other such docu- ous to a person skilled
	e actual completion of theinternational search	Date of mailing of the international se	
	2 September 1998	1 1 09 195	18
Name and	European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswyk	Authorized officer	
	Tel (+31-70) 340-2040. Tx: 31 651 epo nl. Fax: (+31-70) 340-3016	Chambonnet, F	

In tional Application No PCT/US 97/23819

C.(Continu	Relevant to claim No	
Category	Citation of document, with indication where appropriate of the relevant passages	
Y	LINK. C.D.: "Transgenic Caenorhabditis elegans as model system to study amyloid formation and toxicity" NEUROBIOLOGY OF AGING. vol. 15, no. suppl. 1, 29 July 1994, page S33.XP002065640 see the whole document	1.5-7
X	WO 95 11968 A (ATHENA NEUROSCIENCES INC :LILLY CO ELI (US)) 4 May 1995 see page 38. line 3 - page 39. line 22	11.12,19
Y	WO 93 23533 A (TSI CORP) 25 November 1993 see page 27, paragraph 3 - page 28. paragraph 2	11-15
Y	CHALFIE M ET AL: "GREEN FLUORESCENT PROTEIN AS A MARKER FOR GENE EXPRESSION" SCIENCE, vol. 263. no. 5148. 11 February 1994. pages 802-805. XP002003599 see the whole document	1.2.6,7. 9.11. 13-17. 20.21
Y	MOSS, J. B. ET AL.: "Green fluorescent protein marks skeletal muscle in murine cell lines and zebrafish" GENE. vol. 173, no. 1, 1996. pages 89-98. XP002065641 AMSTERDAM NL see the whole document	1,2,9. 13,16,17
	•	

Int. Ational application No. PCT/US 97/23819

Box 1 Observations where certain claims we	ere found unsearchable (Continuation of item 1 of first sheet)
This International Search Report has not been establis	shed in respect of certain claims under Article 17(2)(a) for the following reasons:
Claims Nos.: because they relate to subject matter not req	uired to be searched by this Authority, namely:
	
Claims Nos.: because they relate to parts of the International Someon extent that no meaningful International Someon extent that no meaningful International Someon extent that no meaningful International Some	inal Application that do not comply with the prescribed requirements to such earch can be carned out, specifically:
	·
3. Claims Nos.: because they are dependent claims and are	e not drafted in accordance with the second and third sentences of Rule 6 4(a).
Box II Observations where unity of inventi	on is lacking (Continuation of item 2 of first sheet)
	inventions in this international application, as follows:
see additi	onal sheet
As all required additional search fees were searchable claims.	timely paid by the applicant, this International Search Report covers all
As all searchable claims could be searche of any additional fee.	d without effort justifying an additional fee, this Authority did not invite payment
3. As only some of the required additional se covers only those claims for which fees we	parch fees were timely paid by the applicant, this International Search Report ere paid, specifically claims Nos.:
No required additional search fees were to restricted to the invention first mentioned	timely paid by the applicant. Consequently, this International Search Report is in the claims; it is covered by claims Nos.:
Remark on Protest	The additional search fees were accompanied by the applicant's protest.
	No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

and partially 11-23 1. Claims: 1-10

a transparent animal comprising:

(a) a toxic transgene:

- (b) an inducible promotor operably linked to a reporter gene; wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo; and a method for providing such a dually transgenic transparent non-human animal comprising:
 - a) providing:
 - i) a first and a second non-human transparent animals;
 - ii) a first toxic transgene; and
 - iii) a second transgene, comprising an inducible promotor operably linked to a reporter gene, wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo,
 - b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and
 - c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic transparent non-human animal wherein said toxic transgene is expressed and induces the promotor so that the said reporter gene is also expressed and detectable in vivo;
 - 2. Claims: partially 11-23

a method to produce a dually transgenic non-human animal with a toxic transgene and a reporter are expressed comprising:

- a) providing:
- i) a first and a second non-human animal;
- ii) a first toxic transgene; and
- iii) a second transgene, comprising a reporter,
- b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic non-human animal wherein said toxic transgene and said reporter are expressed

in the cases non relevant of the first described subject, that means where it is not a transparent animal comprising :

(a) a toxic transgene;

(b) an inducible promotor operably linked to a reporter gene; wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo.

Information on patent family members

Ir. attorial Application No
PCT/US 97/23819

Patent document cited in search report		Publication date	Patent lamily member(s)	Publication date
CA 2088379	Α	30-07-1994	NONE	
wo 9511968	A	04-05-1995	US 5604102 A US 5612486 A AU 8079894 A AU 8080994 A CA 2174429 A CA 2174632 A EP 0736106 A EP 0730643 A JP 9508196 T JP 9507746 T WO 9511994 A	18-02-1997 18-03-1997 22-05-1995 22-05-1995 04-05-1995 04-05-1996 11-09-1996 19-08-1997 12-08-1997 04-05-1995
W0 9323533	Α	25-11-1993	AU 4240293 A	13-12-1993